Inverse Modeling to Verify California's Greenhouse Gas Emission Inventory

Report Version 1.3

Contract Number 09-348

Investigators:

Marc L. Fischer and Seongeun Jeong
California State University East-Bay
and Lawrence Berkeley National Laboratory

April, 2013

Prepared for the
California Air Resources Board and
the California Environmental Protection Agency

DISCLAIMER

The statements and conclusions in this Report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in the connection with material reported herein is not to be construed as actual or implied endorsement of such products.

ACKNOWLEDGEMENT

This report was submitted in fulfillment of California Air Resources Board (CARB) Contract number 09-348. We thank Arlyn E. Andrews, Laura Bianco, James M. Wilczak, Dave Field, Dave Bush, Edward Wahl, Ken Reichl, Fabien Guerin, Yuchen Yi, and particularly Jon Kofler for assistance with measurements at Walnut Grove, California (WGC) and analysis of data from radar wind profiler sites, John Lin, Christoph Gerbig, Steve Wofsy, Janusz Eluszkiewicz, Thomas Nehrkorn for sharing the Stochastic Time-Inverted Lagrangian Transport (STILT) code and providing advice, Chris Potter and William Salas for sharing modeled methane (CH_4) emission for use as a priori estimates, Ed Dlugokencky and Colm Sweeney for sharing data for CH_4 background estimates, Larry Hunsaker, Marc Vayssières, Joseph Fischer, and Webster Tassat for sharing CARB CH_4 emissions information, Ying-Kuang Hsu and Patrick Vaca for the operation and data processing of CARB greenhouse gas monitoring network, and Krishna Muriki for assistance running the Weather Research and Forecasting (WRF) and STILT models on the LBNL-Lawrencium computer cluster. We gratefully acknowledge the National Oceanic and Atmospheric Administration (NOAA) Air Resources Laboratory (ARL) for the use of the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model underlying STILT, and the National Centers for Environmental Prediction (NCEP) for the provision of the North American Regional Reanalysis (NARR) meteorology. We also thank Jean Bogner, Eric Crosson, Guido Franco, Eileen McCauley, and Tony VanCuren for valuable comments. This work was predominantly supported by CARB, with additional support from the California Energy Commission, and the US Department of Energy.

TABLE OF CONTENTS

DISCLAIMER i
ACKNOWLEDGEMENT ii
TABLE OF CONTENTS iii
LIST OF FIGURES iv
LIST OF TABLES vi
ABSTRACT vii
1. Introduction 10
2. Approach 10
2.1. CH_4 Measurements and Boundary Conditions 12
2.2. A priori CH_4 Emission Maps 15
2.3. Atmospheric Transport Modeling 21
2.4. Bayesian Inverse Model 24
2.5. Uncertainty Analysis 26
3. Results and Discussion 30
3.1. CH_4 Mixing Ratio 30
3.2. Footprints 34
3.3. Bayesian Inverse Analysis 36
3.3.1. Linear Analysis 36
3.3.2. Bayesian Region Analysis 38
3.3.3. Bayesian Source Analysis 41
4. Conclusions and Recommendations 45
5. References 47

LIST OF FIGURES

Figure 1. Inverse modeling approach used in the study. 12

Figure 2. 3-hourly signal comparison: all measured CH_4 signal (gray open circle), measured CH_4 signal during noon - afternoon hours (black filled circle), WRF-STILT predicted CH_4 signal + WRF-STILT predicted CH_4 background signal during noon – afternoon hours (blue open circle), and WRF-STILT predicted CH_4 background signal using the 3-D curtain (red dots). 14

Figure 3. California-specific CH_4 emission (nmol m^{-2} s^{-1}) maps (available at (http://calgem.lbl.gov/prior_emission.html) for source sectors: (a) LF, (b) WW, (c) DLS, (d) NDLS, (e) NG, and (f) PL. 16

Figure 4. California-specific CH_4 emission maps for source sectors that have seasonal components: (a) CP and (b) WL. 17

Figure 5. Sub-region classification for the inverse analysis. Emissions from these regions are adjusted by the corresponding scaling factors estimated by the inverse model. 18

Figure 6. (a) California-specific total CH_4 emissions (nmol m^{-2} s^{-1}), and (b) EDGAR42 CH_4 total emissions (nmol m^{-2} s^{-1}). 19

Figure 7. Comparison between the California-specific and EDGAR42 emissions by region. ... 20

Figure 8. Schematic diagram showing simulations of backward (opposite to the wind direction) particle trajectories from a tower that are used for footprint calculations. Footprints at the grid cells near the receptor are strong because more particles pass over those grid cells before they are advected and dispersed into different places. The particles above ½ PBLH (colored in gray) do not contribute to the footprint strength. 22

Figure 9. WRF modeling domain configuration with three-level nested domains (d01, d02, and d03 featuring 36, 12, and 4 km resolution, respectively). 24

Figure 10. Location of GHG measurement sites (black) and wind profiler sites (red) in the Central Valley with predicted monthly mean PBL heights (m) for June 2011, 14:00 local standard time (LST) shown in color. 29

Figure 11. Comparison of measured and predicted *Zi* during the month of June for (a) CCO, (b) CCL, (c) LHS, and (d) SAC. For CCO, data from June 2011 data are used while the other sites use data from June 2010. For this summer month, the 5-L LSM scheme was used for all sites. In terms of the PBL scheme, the MYJ scheme was used for all sites except for LHS where the YSU scheme was used. 30

Figure 12. Time series of measured and predicted mixing ratios at the five network sites during September 2010 – June 2011. The measurements are shown for both day and night times

while the predictions are shown for only noon-afternoon well-mixed periods. The prediction was made based on the California-specific emission maps. 32

Figure 13. 3-hourly time series of measured and predicted mixing ratios at the five network sites during noon-afternoon well-mixed periods. The prediction was made based on the California-specific emission maps. 33

Figure 14. Averaged footprints during the noon-afternoon hours for (a) the WGC site and (b) all five sites during May – June 2011. 35

Figure 15. Seasonal mean footprints during the noon-afternoon hours for (a) September – October 2010, (b) November – December 2010, (c) January – February 2011 and (d) March – April 2011. 36

Figure 16. Comparison of California-specific predicted vs. measured CH_4 signals during May 2011 before (left) and after (right) inverse optimization. The light blue circles indicate those removed after the first inversion. 37

Figure 17. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) by region and season based on the California-specific emission model. Only regions with significant emissions are shown. The annual mean prior (gray bar) represents the annual average of seasonally varying emissions and is compared with posterior seasonal emissions (color bars). SO, ND, JF, MA, and MJ denote September-October, November-December, January-February, March-April, and May-June seasons, respectively. 39

Figure 18. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) by region and season based on the EDGAR42 emission model. Only regions with significant emissions are shown. 40

Figure 19. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) for California by source and season based on the California-specific emission model. WW, LF, DLS, NDLS, NG, PL, WL and CP represent wastewater, landfill, dairy livestock, non-dairy livestock, natural gas, petroleum, wetland, and crop agriculture sources, respectively. 42

Figure 20. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) for California by source and season based on the EDGAR42 emission model. AS, EF, GPD, MM, OPR, RT, SW, and WW represent agricultural soils, enteric fermentation, gas production and distribution, manure management, oil production and refineries, road transportation, solid waste, and wastewater, respectively. 43

LIST OF TABLES

Table 1. Measurement Sites and Periods 13

Table 2. Annual Average California-specific CH_4 Emissions by Region and Sector (Tg CO_2eq) 17

Table 3. Regions Used for Inverse Analysis and Corresponding California Counties and Air Basins 18

Table 4. Comparison of CH_4 Emissions by Source between CARB 2008 Inventory and EDGAR42 Emission Model (100-year GWP = 21 g CO_2eq / g CH_4) 21

Table 5. Linear Analysis Results Before and After Bayesian Region Inversion 38

Table 6. Comparison of Annual Posterior CH_4 Emissions (Tg CO_2eq) between the EDGAR42 and California-specific Emission Models Based on Bayesian Region Analysis 41

Table 7. Annual Posterior CH_4 Emissions (Tg CO_2eq) for California by Source Based on the California-specific Emission Model 42

Table 8. Annual Posterior CH_4 Emissions (Tg CO_2eq) for California by Source Based on the EDGAR42 Emission Model 44

Table 9. Summary of Estimated Annual CH_4 Emissions (Tg CO_2eq; 100-year GWP = 21 g CO_2eq / g CH_4) for California 45

ABSTRACT

We estimate regionally resolved methane (CH_4) emissions for California by comparing CH_4 mixing ratios measured at a network of measurement sites in the Central Valley with transport model predictions based on two independent emission maps: a 0.1 degree seasonally varying "California-specific" emission map, calibrated to state-wide by CH_4 emission totals, and the 0.1 degree global EDGAR42 CH_4 emission map. Atmospheric particle trajectories and surface footprints (sensitivity of CH_4 signals to surface emissions) are computed using the Weather Research and Forecasting (WRF) and Stochastic Time-Inverted Lagrangian Transport (STILT) models. Uncertainties due to wind velocity and boundary layer mixing depth are evaluated using measurements from radar wind profilers. Bayesian region analyses of data from the tower network constrains annual average CH_4 emissions from California's Central Valley to between 31.43±2.07 and 28.27±2.00 Tg CO_2eq (assuming a global warming potential of 21 Tg CO_2eq/ Tg CH_4) for the California-specific and EDGAR42 emission models respectively, showing consistency between the two independent models. Extrapolating results to annually averaged CH_4 emissions across all of California totals 1.44±0.15 and 1.94±0.28 times larger than the current inventory estimate for State annual total CH_4 emissions (32 Tg CO_2eq) for the California-specific CH_4 and EDGAR42 CH_4 emission maps, respectively. When emissions from large urban areas are estimated based on a recent study in the larger Los Angeles metropolitan region to better constrain urban emissions, State total CH_4 emissions are estimated to be 1.30 – 1.74 times larger than the current State total CH_4 emissions. These results based on the multiple emission models suggest that the California total of CH_4 emissions would account for approximately 8% - 13% of the State's total greenhouse gas (GHG) emissions, which is significantly higher than the CARB inventory (~6% of total GHG emissions). Spatial resolution of emissions within the influence region reveal seasonality expected from several biogenic sources, including rice agriculture. We expect that additional tower measurements in urban regions (e.g., South Coast Air Basin) will provide the data necessary for a complete analysis of California's CH_4 budget.

EXECUTIVE SUMMARY

Background

Methane (CH_4) is the second highest contributor to climate change among greenhouse gases (GHGs) behind carbon dioxide (CO_2), based on its concentration changes in the atmosphere since the start of the industrial revolution and its ability to absorb infrared radiation. At the regional scale, California currently emits approximately 500 Tg (1 Tg = 1 million metric ton) of CO_2 equivalent (CO_2eq) GHGs, with CH_4 currently estimated to contribute approximately 6% of the total [California Air Resources Board (CARB), 2011].

Methods

This report quantifies regional CH_4 emissions within California with a Bayesian inverse modeling approach, representing the first analysis of CH_4 emissions across a large swath of California and across different seasons using atmospheric observations from multiple sites. The inverse modeling approach follows the approach taken in Zhao et al. [2009], Jeong et al. [2012a] and Jeong et al. [2012b]. We calculate predicted CH_4 signals using two relatively high resolution (0.1 degree) emissions models, a California-specific model and the EDGAR4.2 global emission model. Central to this approach, we quantify model-measurement uncertainties by estimating errors in transport variables (e.g., wind velocities and planetary boundary layer depth) that affect footprints (sensitivity of CH_4 signals to surface emissions in units of concentration/flux) and propagating those errors to produce uncertainty in predicted CH_4 signals. The Bayesian inverse analysis then estimates posterior (optimized) CH_4 emissions for regions (region analysis) and source sectors (source analysis).

Results

This study shows that actual CH_4 emissions based on inverse region analyses are 1.44±0.15 - 1.94±0.28 times larger than the current inventory estimates (32 Tg CO_2eq yr^{-1}). When emissions from large urban areas (e.g., Southern California region) are estimated based on a recent study [Wennberg et al., 2012] in the larger Los Angeles metropolitan region and combined with the emissions from the Central Valley and other non-urban regions, State total CH_4 emissions are 1.30 – 1.74 times larger than the current State total CH_4 emissions. A Bayesian region analysis suggests that the relatively large range of total emissions reflects a current limitation to uniquely resolve urban versus rural CH_4 emissions, particularly from Southern California. A similar source sector analysis suggests that the dominant CH_4 emissions are derived from livestock and landfills, though as with the region analysis, significant differences are obtained with the California-specific and EDGAR42 prior emission maps.

Conclusions

Atmospheric CH_4 measurements can be combined to estimate total CH_4 emissions at regional scales using the inverse modeling approach. Combining results from these tower measurements with results from urban area, California's CH_4 emissions are estimated to be 1.30 – 1.74 times larger than the current State total CH_4 emissions.

Recommendations for work that will likely reduce these uncertainties include:

- While CH_4 emissions from California's Central Valley are well constrained by the current network of measurement sites, emissions from coastal (and predominantly urban) regions remain uncertain. We suggest additional tower measurements in the San Francisco Bay and Southern California areas will be effective in constraining those emissions.

- Data from the current CH_4 measurement network are effective for use in constraining emissions from different regions of California's Central Valley, but cannot be used to uniquely attribute emissions to specific source sectors (e.g., landfills, livestock, petroleum and natural gas). We expect measurements of additional source specific tracers (e.g., VOCs, CO, and potentially CH_4 isotopes) help separate different sources of methane.

- Remaining uncertainty in the inverse model estimates of CH_4 emissions for regions containing measurement sites are dominated by the combination of uncertainty in meteorological modeling of trace gas transport (in winter) and estimation of background signals (in summer). We suggest that further work on meteorological modeling and background estimation will be effective in identifying and reducing these sources of uncertainty.

PROJECT REPORT

1. Introduction

Methane (CH_4) is the second highest contributor to climate change among greenhouse gas (GHG) behind carbon dioxide (CO_2), based on its concentration changes in the atmosphere since the start of the industrial revolution, the long residence time of CH_4 and its ability to absorb infrared radiation. Earth's CH_4 has increased by about 150% since 1750 in concentration, and accounts for ~ 25% of the global total radiative forcing from all long-lived and globally mixed GHGs [Hofman et al., 2006; Montzka et al., 2011]. Given the importance of CH_4 as a GHG it is important to be able to quantify changes in emissions. However, there exists a large uncertainty in bottom-up emission inventory models that take known natural and anthropogenic sources of CH_4 to produce emission estimates due to lack of understanding of emission processes and driving data. Mathematical inversion models, which use concentration changes in CH_4 and transport to infer sources, provide an effective tool for understanding CH_4 emissions. Correspondingly, attention has focused on inverse model assessment of global [Gimson and Uliasz, 2003; Houweling et al., 1999], and regional [Kort et al., 2008; Zhao et al., 2009; Jeong et al., 2012a] CH_4 sources.

At the regional scale, California currently emits approximately 500 Tg (1 Tg = 1 million metric ton) of CO_2 equivalent GHGs, with CH_4 currently estimated to contribute approximately 6% of the total [California Air Resources Board (CARB), 2011]. Because California has committed to an ambitious plan to reduce GHG emissions to 1990 levels by 2020 through Assembly Bill 32 (AB-32), verifying the success of control strategies will require accounting for CH_4 emissions.

This report quantifies regional CH_4 emissions from California within a Bayesian inverse modeling framework, representing the first analysis of CH_4 emissions in California using atmospheric observations from multiple sites across different seasons during 2010 - 2011. The work expands on studies by Zhao et al. [2009] and Jeong et al. [2012a] that quantified CH_4 emissions from central California using a single tower near Walnut Grove, California (WGC). In Section 2, we describe the methods we employed, including atmospheric measurements, *a priori* CH_4 emissions inventories, mesoscale meteorology and trajectory transport modeling, and the Bayesian inverse method. Section 3 describes results, including the seasonal variations in calculated footprints, and the inferred surface emissions of CH_4 from California for different regions and sources based on simple correlation analysis and a Bayesian inverse analysis. Section 4 summarizes the results and presents the recommendations for CH_4 inverse modeling at the regional scale, highlighting the importance of uncertainty in the spatial distribution of *a priori* emissions, and the value of multiple measurement stations.

2. Approach

The inverse modeling framework employed in this multi-site study builds on the approach taken in Zhao et al. [2009], Jeong et al. [2012a] and Jeong et al. [2012b] where GHG measurements from a single tower were used. The inverse modeling approach used in this study is illustrated in Figure 1. As shown in the figure, the Bayesian inverse model requires two direct inputs (two

arrows are directed to the "Bayesian Inversion" box): 1) CH_4 tower measurements (measured signals), and 2) predicted signals. By comparing measured signals with predicted signals, the Bayesian inverse model estimates scaling factors for surface emissions (i.e., spatially/temporally resolved emission inventory) such that the scaled surface emissions yield predicted signals that are statistically consistent with measurements. Predicted signals are calculated as a linear combination of 1) footprints, which represent the sensitivity of signals measured at different sites to emissions across the landscape, and 2) surface emissions (see Section 2.3 for details on footprint calculations).

Because predicted signals represent local enhancements of emissions, upwind background signals entering the study domain are needed to compare predicted signals with measured signals, which include both local and background signals. CH_4 boundary values are the initial CH_4 mixing ratios that represent the upstream concentrations at the western domain boundary (130°W) and often called CH_4 background concentrations (see Section 2.1 for boundary values). Footprints are quantified using a Lagrangian model for air parcels arriving at a tower. Such footprints link the observed concentrations at a specific location and height (e.g., measurement tower) to surface fluxes within a large area. Numerical meteorological model outputs are used to define paths traveled by parcels of air, or trajectories, which are a basis for footprint estimates. In this study we use the coupled WRF-STILT model for trajectory calculations. The WRF-STILT model has been used to constrain GHG emissions in several studies including airborne measurement-based (e.g., Gerbig et al., 2003; Kort et al., 2008) and tower measurement-based inversions (e.g., Zhao et al., 2009; Miller et al., 2012; Jeong et al., 2012a; Jeong et al., 2012b). Errors in modeling footprints due to uncertainties in winds and planetary boundary layer (PBL) heights contribute to uncertainties in inversion results and confidence levels associated with optimized emissions values. The result of the Bayesian inverse model is a set of optimized scaling factors for region or source emissions (see Section 2.4 for details on Bayesian inversion).

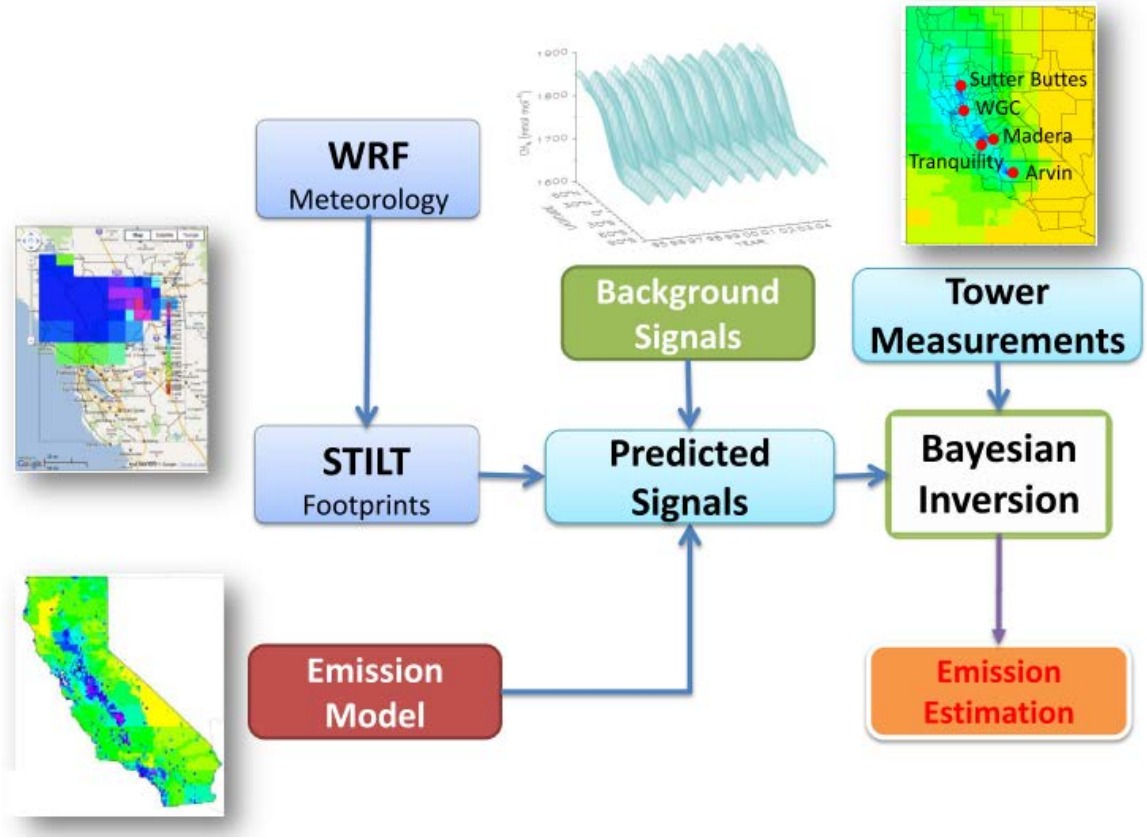

Figure 1. Inverse modeling approach used in the study.

2.1. CH_4 Measurements and Boundary Conditions

CH_4 measurements were made at the collaborative five-site GHG network in California's Central Valley. In addition to the Central Valley sites, CH_4 measurements were also made at Mt. Wilson, but these measurements are not employed in this study because we have not developed a well-tested meteorological model for atmospheric transport at Mt. Wilson at this time. Table 1 summarizes the information for the measurement sites in the measurement network and the measurement periods used in the inverse analysis. The Arvin (ARV) site is located at the southern end of the San Joaquin Valley and constrains emission sources from livestock, and gas and petroleum production fields. The Madera (MAD) and Tranquility (TRA) sites are located in the center of the San Joaquin Valley, constraining emission sources mainly from livestock. The Sutter Buttes (STB) site located in the Sacramento Valley represents an emission region that has dominant CH_4 emissions from rice agriculture. As described in detail in Zhao et al. [2009] and Jeong et al. [2012a], the Walnut Grove (WGC) site represents an emission region with mixed CH_4 emission sources such as crop agriculture, livestock, natural gas fields, wetlands and urban emissions.

CH_4 measurements at WGC were made at 91 and 483 m above ground level on a tall tower, beginning in September 2007. The CH_4 mixing ratios at each height are measured every 15 minutes and averaged into the 3-hour means used in this study. CH_4 mixing ratio is defined as

the amount (usually in mole) of CH_4 divided by the total amount of an air mixture and expressed as (nano mol) / mol or parts per billion (ppb). As in Zhao et al. [2009] and Jeong et al. [2012a], CH_4 measurements at 91 m are used for inverse modeling. Detailed information about these measurements is described by Zhao et al. [2009] and Jeong et al. [2012a]. All other stations are measured at 10 meters above the ground using the same type of instruments and calibrated with standard gases from NOAA every six months. Each instrument is programmed to measure from precision check standard gases every 11 hours to ensure data quality. After examining precision checks and removing special events (e.g., changing filters), raw data collected every few seconds are averaged into 3-hourly measurements for inverse modeling.

Table 1. Measurement Sites and Periods

Site Name	Height[a]	Elevation[b]	Latitude	Longitude	Measurement Period (yyyymm)
Arvin (ARV)	10 m	158 m	35.24°N	118.79°W	201009 - 201106
Madera (MAD)	10 m	81 m	36.87°N	120.01°W	201009 - 201106
Sutter Buttes (STB)	10 m	640 m	39.21°N	121.82°W	201105 - 201106
Tranquility (TRA)	10 m	59 m	36.63°N	120.38°W	201009 - 201106
Walnut Grove (WGC)	91 m	0 m	38.27°N	121.49°W	201009 - 201106

[a]Above ground level (a.g.l.)
[b]Above sea level (a.s.l.)

As described previously, the measured signals at each measurement site include both CH_4 enhancements and upwind boundary conditions while the WRF-STILT model simulates only local CH_4 enhancements. Therefore, upstream CH_4 boundary conditions need to be estimated to compare local measured CH_4 signals with predicted CH_4 in inverse modeling. The CH_4 upstream boundary product used in this study is similar to the one used in Jeong et al., [2012b]. CH_4 boundary values were estimated using data from the Pacific coast aircraft network CH_4 profiles (http://www.esrl.noaa.gov/gmd/ccgg/aircraft/) and remote Pacific marine boundary layer sampling sites (http://www.esrl.noaa.gov/gmd/ccgg/flask.html) within the NOAA Earth System Research Laboratory (ESRL) Cooperative Air Sampling Network. The data were smoothed and interpolated to create a three-dimensional (3-D) curtain, varying with latitude, height and time. Therefore, this 3-D background curtain provides estimates for CH_4 background concentrations at a given latitude (at an upwind longitude), elevation, and time. As in Zhao et al. [2009] and Jeong et al. [2012a], predicted background values are computed for each footprint simulation by sampling the curtain at each of the 500 STILT trajectory endpoints (near 130°W) and calculating the average value.

Figure 2 shows the 3-hourly measured signal, background signal and predicted signal using the California-specific *a priori* CH_4 emissions for the five network sites. Predicted signals are shown only for the well-mixed periods (noon – afternoon). For inverse analysis, the hourly measurements and predicted background signals are aggregated into 3-hourly time periods as in Jeong et al. [2012a]. Unlike WGC, the other sites do not have multiple measurement levels. Therefore, it is difficult to identify well-mixed periods without using vertical CH_4 gradients from multiple-level measurements. For inverse analyses, we use data during 12 – 17 hours (local standard time or LST) except for winter (11 – 16 LST) during which we found that WRF-

simulated PBL heights tend to collapse earlier compared to wind profiler measurements. As can be seen in the figure, most of the high mixing ratios are observed during the night time when the PBL height is low. In general, the variability in measured CH_4 is larger in winter than the other seasons for most of the sites. The minimum measured signals approximate the predicted background CH_4, suggesting that the estimated background signals are reasonable and there is no significant bias in the measured signals.

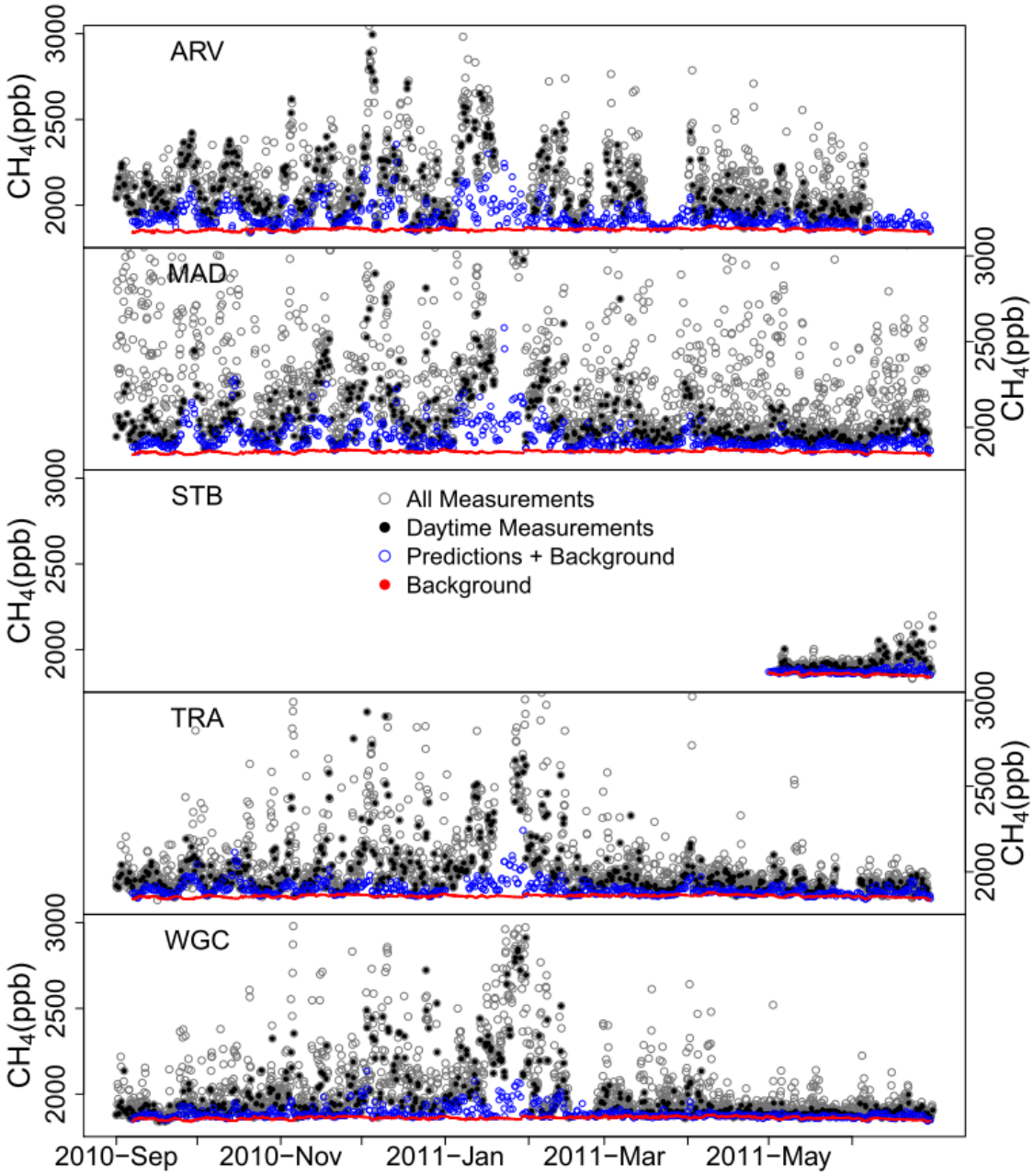

Figure 2. 3-hourly signal comparison: all measured CH_4 signal (gray open circle), measured CH_4 signal during noon - afternoon hours (black filled circle), WRF-STILT predicted CH_4 signal + WRF-STILT predicted CH_4 background signal during noon – afternoon hours (blue open circle), and WRF-STILT predicted CH_4 background signal using the 3-D curtain (red dots).

2.2. A priori CH_4 Emission Maps

This work adopts the *a priori* CH_4 emission model from Jeong et al. [2012a], which provides a high spatial resolution (0.1º × 0.1º) for California and has seasonal components for wetlands and crop agriculture. As described in Jeong et al. [2012a], the high resolution emission model was prepared by scaling to the CARB 2008 inventory by sector [CARB, 2010]. The considered sectors include: crop agriculture (CP), landfills (LF), dairy livestock (DLS), non-dairy livestock (NDLS), natural gas (NG), petroleum (PL), wastewater (WW), and wetlands (WL). Figure 3 shows California-specific CH_4 emission maps for sectors without temporal variation.

Because there is no specific emission estimate for wetlands from CARB, wetland CH_4 emissions are taken from monthly averages of the Carnegie-Ames-Stanford-Approach CH_4 (CASA-CH_4) model from Potter et al. [2006]. Also, seasonally varying CH_4 emissions for CP CH_4 sources were taken from the denitrification and decomposition model (DNDC) output (assuming the 1983, high irrigation case) described by Salas et al. [2006]. Monthly averaged CH_4 emission maps for county level agricultural CH_4 fluxes are used. The temporally-varying emission maps for CP and WL are averaged annually and shown in Figure 4.

Table 2 summarizes the annual mean CH_4 emissions for the California-specific emission model, which is calibrated to the CARB 2008 inventory [CARB, 2010], by region and sector. In this study, CH_4 emissions are scaled to CO_2 equivalent using a 100-year global warming potential (GWP) of 21 g CO_2 eq / g CH_4 [IPCC, 1995]. The sub-regions (shown in Figure 5) in Table 2 were classified by considering the emission sources and measurement sites, and roughly follow the California Air Basins. Large California Air Basins such as the Sacramento Valley Basin and the San Joaquin Valley Basin were divided into two basins. For example, the Sacramento Valley Basin was divided into Region 6 and Region 11, considering emissions sources; emissions from rice agriculture are concentrated in Region 6 while rice agriculture emissions in Region 11 are small. On the other hand, due to weak footprint influence from the measurement sites located in the Central Valley, we combined the basins (Mojave Desert, South Coast, San Diego, Salton Sea) in the Southern California region into Region 10. Regions used for inverse analysis and corresponding California counties and air basins are shown in Table 3.

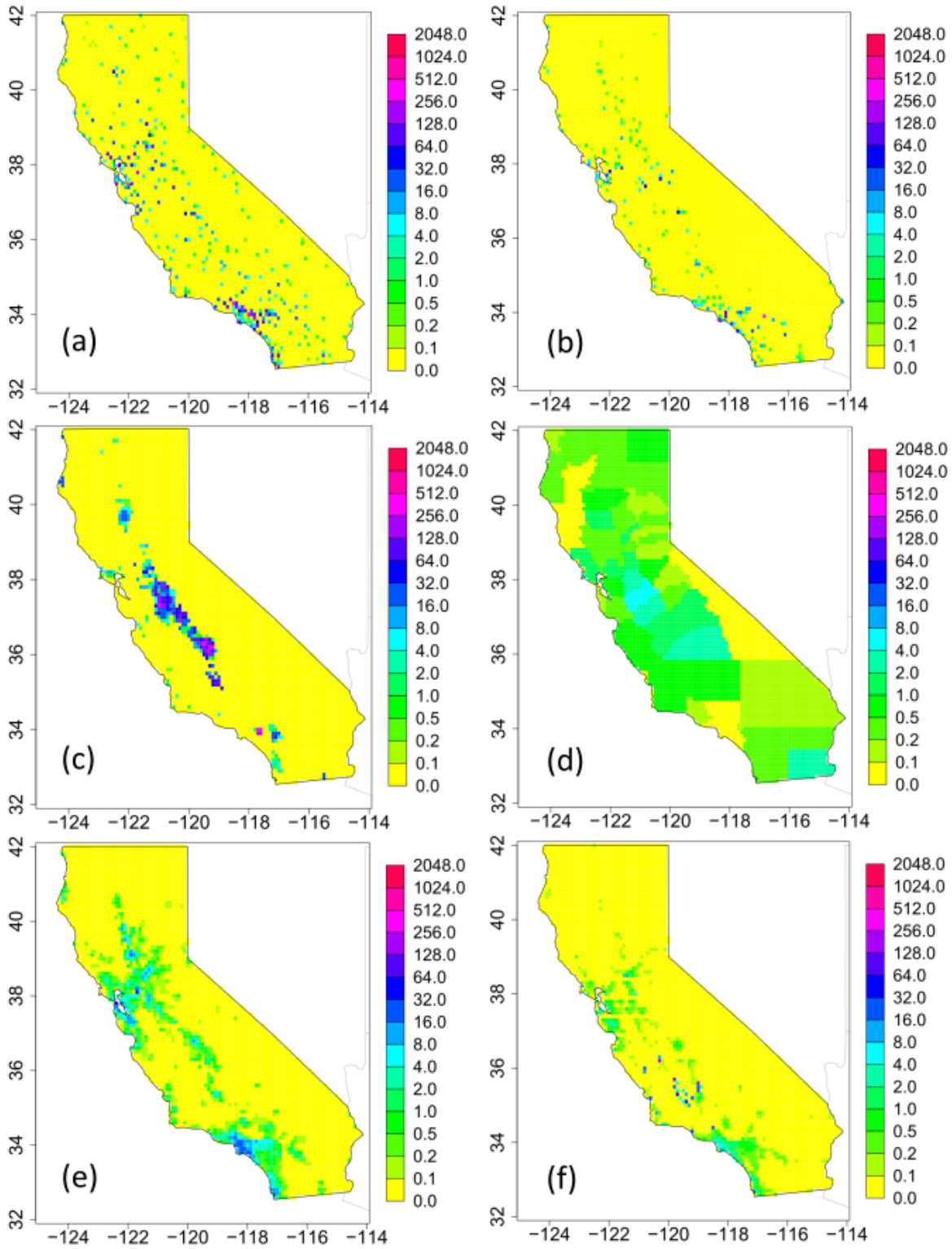

Figure 3. California-specific CH_4 emission (nmol m^{-2} s^{-1}) maps (available at (http://calgem.lbl.gov/prior_emission.html) for source sectors: (a) LF, (b) WW, (c) DLS, (d) NDLS, (e) NG, and (f) PL.

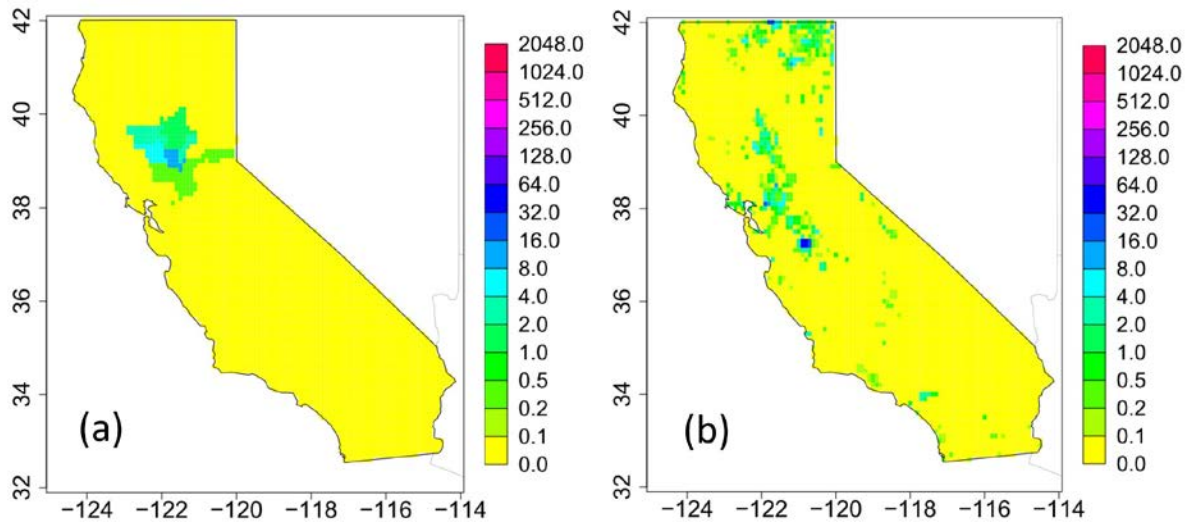

Figure 4. California-specific CH_4 emission maps for source sectors that have seasonal components: (a) CP and (b) WL.

Table 2. Annual Average California-specific CH_4 Emissions by Region and Sector (Tg CO_2eq)

Sector/ Region[a]	R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	Sector Total
CP	0.00	0.00	0.00	0.01	0.00	0.50	0.00	0.01	0.00	0.00	0.01	0.00	0.02	**0.54**
LF	0.02	0.04	0.11	0.08	0.03	0.46	0.87	0.19	0.34	4.00	0.10	0.29	0.06	**6.60**
DLS	0.00	0.00	0.01	0.10	0.01	0.36	0.08	3.79	0.02	1.71	0.03	5.77	0.01	**11.90**
NDLS	0.03	0.10	0.11	0.06	0.17	0.19	0.12	0.54	0.11	0.64	0.07	1.00	0.03	**3.17**
NG	0.00	0.01	0.04	0.02	0.01	0.33	0.33	0.10	0.05	0.91	0.02	0.11	0.03	**1.95**
PL	0.00	0.00	0.05	0.00	0.00	0.03	0.05	0.02	0.07	0.19	0.00	0.71	0.00	**1.13**
WW	0.00	0.09	0.02	0.01	0.00	0.03	0.17	0.08	0.06	1.33	0.01	0.11	0.01	**1.92**
WL	0.01	0.00	0.00	0.00	0.22	0.18	0.03	0.27	0.01	0.03	0.01	0.02	0.01	**0.79**
Region Total	**0.07**	**0.24**	**0.34**	**0.30**	**0.44**	**2.08**	**1.65**	**4.99**	**0.67**	**8.80**	**0.25**	**8.02**	**0.16**	**28.00**

[a]See Figure 5 for region classification in a map

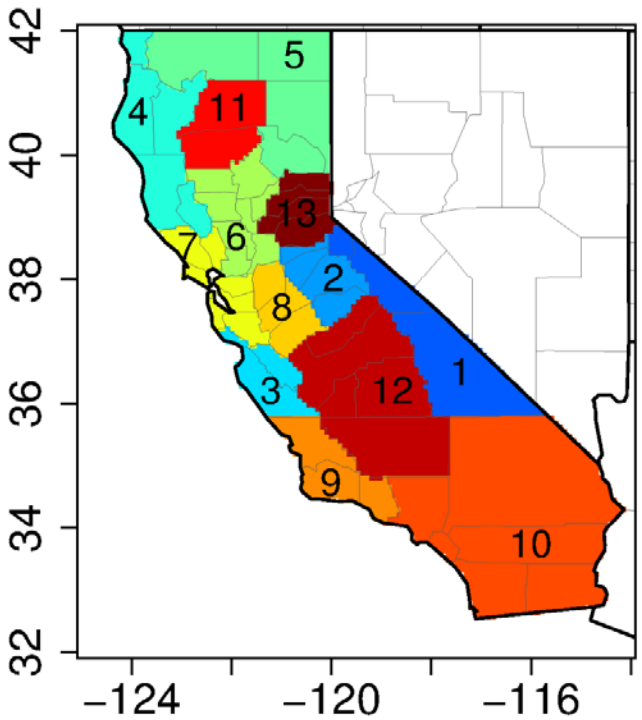

Figure 5. Sub-region classification for the inverse analysis. Emissions from these regions are adjusted by the corresponding scaling factors estimated by the inverse model.

Table 3. Regions Used for Inverse Analysis and Corresponding California Counties and Air Basins

Regions	Counties	California Air Basins[a]
1	Mono, Inyo, Alpine	Great Basin Valleys
2	Calaveras, Amador, Mariposa, Tuolumne	Mountain Counties
3	Monterey, San Benito, Santa Cruz	North Central Coast
4	Mendocino, Lake, Trinity, Del Norte, Humboldt	North Coast
5	Lassen, Siskiyou, Plumas, Modoc	Northeast Plateau
6	Colusa, Glenn, Butte, Yolo, Yuba, Sacramento, Solano, Sutter	Sacramento Valley
7	Alameda, Napa, Contra Costa, Santa Clara, Sonoma, San Mateo, Marin, San Francisco	San Francisco Bay Area
8	Merced, San Joaquin, Stanislaus	San Joaquin Valley
9	Ventura, Santa Barbara, San Luis Obispo	South Central Coast
10	San Diego, Los Angeles, Orange, San Bernardino, Imperial, Riverside	Mojave Desert, South Coast, San Diego, Salton Sea
11	Tehama, Shasta	Sacramento Valley
12	Fresno, Madera, Kern, Kings, Tulare	San Joaquin Valley
13	Sierra, Nevada, Placer, El Dorado	Lake Tahoe, Mountain Counties

[a]Some basins are divided or combined to represent an emission region for inversion

The EDGAR42 (European Commission Joint Research Centre (JRC) and Netherlands Environmental Assessment Agency, Emission Database for Global Atmospheric Research (EDGAR), release version 4.2, 2011, http://edgar.jrc.ec.europa.eu) CH_4 emission model also provides high-resolution emission maps and its estimates are compared with estimates from the California-specific model (Figure 6). Compared with the California-specific model, EDGAR42 generally shows a similar spatial distribution of CH_4 emissions. However, EDGAR42 shows larger emissions in urban areas than the California-specific model. For the Central Valley, the California-specific model estimates higher emissions than EDGAR42, mainly due to the higher estimates of livestock emissions.

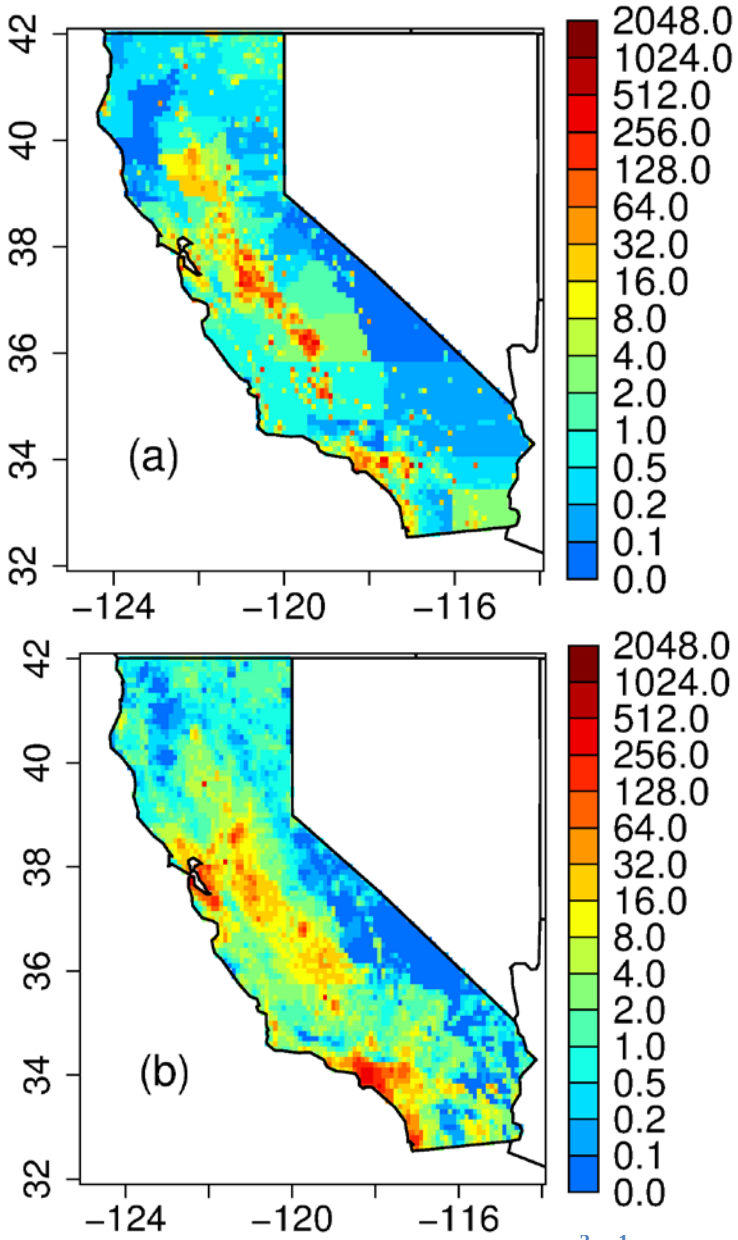

Figure 6. (a) California-specific total CH_4 emissions (nmol m^{-2} s^{-1}), and (b) EDGAR42 CH_4 total emissions (nmol m^{-2} s^{-1}).

Figure 7 shows the comparison between the California-specific and EDGAR42 emissions by region. As can be seen in the figure, EDGAR42 shows more weight in urban areas than the California-specific model. For example, for Regions 7 and 10, which include San Francisco Bay Area and the Southern California Air Basin, respectively, EDGAR estimates significantly higher CH_4 emissions than the California-specific model. Currently, Region 10 is a single large sub-region due to relatively weak sensitivity from the measurement sites in the Central Valley. However, Region 10 needs to be divided into smaller sub-regions when more measurements sites are available in the region. On the other hand, the California-specific model estimates higher emissions than EDGAR in the Central Valley. For example, the California-specific model has higher emissions than EDGAR42 in Region 8 where livestock is a dominant emission source (more than 80%). The annual total emissions for California from the California-specific model and the EDGAR42 model are 28.0 and 38.3 Tg CO_2eq, respectively. Table 4 shows the comparison of CH_4 emissions by source between the CARB 2008 inventory (used to scale the California-specific emissions; CARB, 2010) and the EDGAR42 emission model.

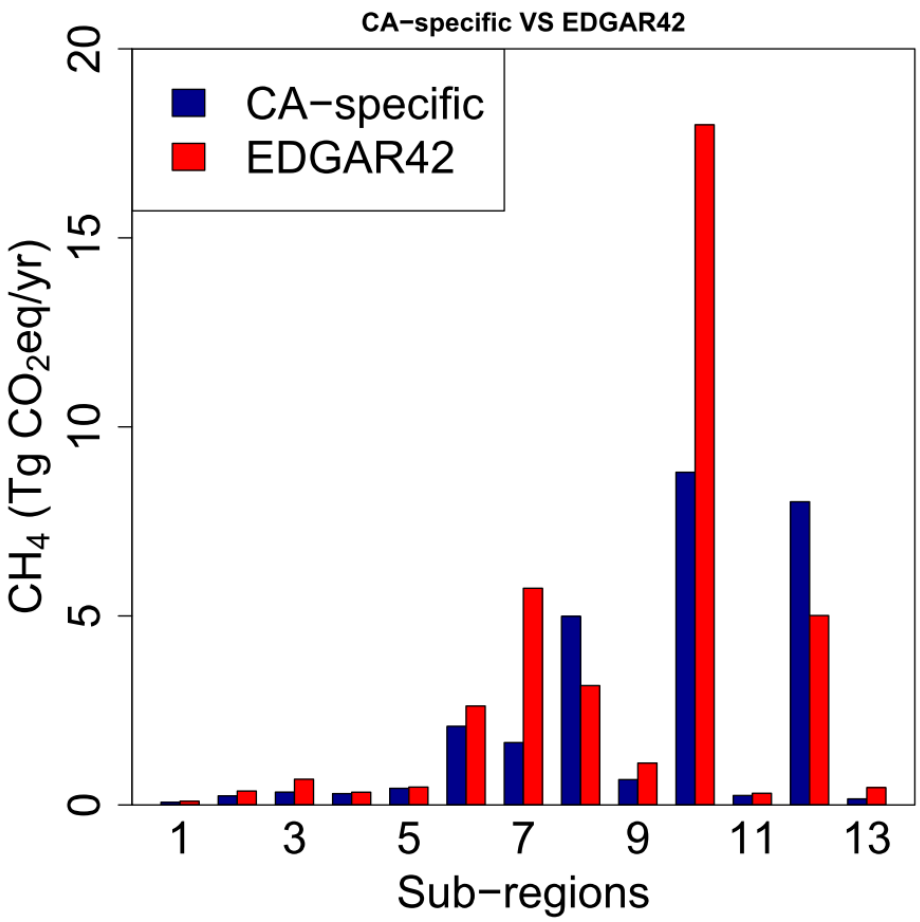

Figure 7. Comparison between the California-specific and EDGAR42 emissions by region.

Table 4. Comparison of CH_4 Emissions by Source between CARB 2008 Inventory and EDGAR42 Emission Model (100-year GWP = 21 g CO_2eq / g CH_4)

CARB Category	CARB Emission (Tg CO_2eq)	EDGAR42 Category	EDGAR42 Emission (Tg CO_2eq)
Rice crop area	0.5	Agricultural soils	0.7
Landfill	6.7	Solid waste	12.7
Dairy cows	11.9	Enteric fermentation	7.3
Non-dairy cows	3.2	Manure management	2.4
Natural gas pipeline	1.9	Gas production and distribution	10.2
Extraction, mobile, refining	1.1	Oil production and refineries	0.5
Wastewater treatment	1.9	Waste water	3.2
Others	1.2		1.4
CARB Total	28.5	EDGAR Total	38.3

2.3. Atmospheric Transport Modeling

Predicted contributions to CH_4 mixing ratios from emissions within the modeling domain are calculated as **FE**, where **F** is footprint strength (an $n \times m$ matrix), and **E** (an $m \times k$ matrix) is the *a priori* CH_4 emissions (see Section 2.4 for details on predicted signal calculation). Footprints represent the sensitivity of the mixing ratio at the receptor location (i.e., measurement site) to surface sources, in units of ppb/(nmol m^{-2} s^{-1}). Footprints are calculated from particle trajectories simulated using the STILT model [Lin et al., 2003, 2004]. Figure 8 below shows particles moving from upwind locations to the measurement site using the STILT model. In the time-reverse sense, the STILT model transports ensembles (e.g., 500 particles) of particles (air parcels) backwards in time (e.g., 7 days) from a receptor point. By releasing particles backwards, we identify the origin of emission sources that contribute to the receptor. In general terms, GHG emission sources upwind produce GHG signals (in mixing ratio units of ppb) at the measurement site as the product of 1) the averaged (e.g., hourly) measurement sensitivity to emissions, or footprint, and 2) the emissions (in flux units of nmol m^{-2} s^{-1}).

As shown in Figure 8, footprints are estimated by counting the time the particle spends in the surface-influenced region, defined as ½ of the planetary boundary layer height (PBLH), and inversely weighted by dilution into the height of the boundary layer. In Figure 8, the dark colored particles within ½ PBLH are assumed to contribute to the footprint, while particles colored in gray above ½ PBLH are not. The footprint for each grid cell (~ 10 km × 10 km) is calculated by counting the number of particles that pass over the grid cell weighted by the amount of time the particles spend in the gird cell. Therefore, if more particles stay in a grid cell over a long period of time, the signal at the measurement site is more sensitive to the grid cell. In this study, 500 particles are released at location $\mathbf{X}_r$ and time t_r, where r represents the receptor. The footprint as function of location and time $f(\mathbf{X}_r, t_r \mid x_i, y_j, t_m)$ links mixing ratios at the receptor $c(\mathbf{X}_r, t_r)$ to discretized surface

emissions $E(x_i, y_j)$ at location (x_i, y_j). Here we assume that CH_4 emissions are constant in time (t_m) and t_m is factored out of *E*. In other words, the footprint converts the surface emissions to a mixing ratio enhancement observed at the measurement site. Therefore, the predicted local signals are ca-lculated by integrating the product of the footprint maps and the *a priori* emission maps over space and time (e.g., 7 days) as

$$c(\mathbf{X}_r, t_r) = \sum_{i,j,m} f(\mathbf{X}_r, t_r \mid x_i, y_j, t_m) \cdot E(x_i, y_j)$$

where the local signal $c(\mathbf{X}_r, t_r)$ is calculated by subtracting the background signal $c_{BG}(\mathbf{X}_r, t_r)$ from the total measured signal $c_{tot}(\mathbf{X}_r, t_r)$.

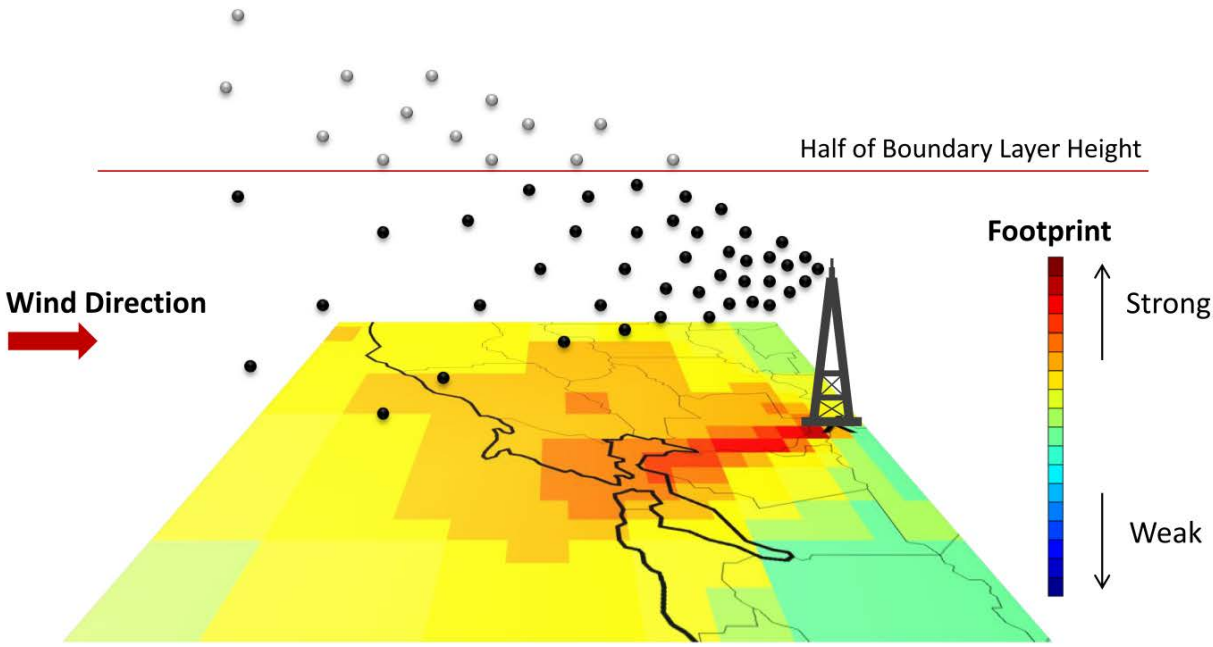

Figure 8. Schematic diagram showing simulations of backward (opposite to the wind direction) particle trajectories from a tower that are used for footprint calculations. Footprints at the grid cells near the receptor are strong because more particles pass over those grid cells before they are advected and dispersed into different places. The particles above ½ PBLH (colored in gray) do not contribute to the footprint strength.

As in Jeong et al. [2012a], 500 particles are released hourly at each measurement site and transported backward in time 7 days to ensure that the majority of the particles reach positions representative of the upwind boundary conditions. This means that predicted signals are calculated for each hour although 3-hourly averaged predictions are used for inverse analysis. We use the coupled WRF-STILT (Weather Research and Forecasting and Stochastic Time-Inverted Lagrangian Transport) model for particle trajectory simulations [Lin et al., 2003; Skamarock et al., 2008; Nehrkorn et al., 2010]. The WRF-STILT model has been used to constrain GHG emissions in many studies including airborne measurement-based (e.g., Gerbig et al., 2003; Kort et al., 2008) and tower measurement-based (e.g., Zhao et al., 2009; Jeong et al.,

2012a; Jeong et al., 2012b) inversions. An ensemble of 500 STILT particles are run backwards in time for 7 days driven with meteorology from the WRF model (version 3.2.1) [Skamarock et al., 2008]. Hourly predicted signals based on WRF-STILT are aggregated into 3-houly averages for inverse modeling.

The WRF model simulations closely follow those described in Jeong et al. [2012a, 2012b] with some modifications, which are summarized here. We use version 3.2.1 of the WRF model [Skamarock et al., 2008] instead of WRF2.2. Five domains (d01 – d05) of 36, 12, 4, and two 1.3 km resolutions were used in the WRF simulations. The 4-km domain (i.e., d03) was configured to represent most of California with the two 1.3-km nested domains (d04 and d05) that cover the San Francisco Bay Area and the metropolitan area of Los Angeles, respectively. In this study, we used the WRF meteorology within the d01, d02 and d03 domains to drive the STILT model because the GHG measurement sites are located in the Central Valley. Figure 9 shows the WRF d01, d02 and d03 domains used in this study. The WRF model was run with two-way nesting instead of one-way nesting used in Jeong et al. [2012a]. As in Jeong et al. [2012a], 50 vertical levels were employed to resolve planetary boundary layer (PBL) heights over complex terrain features of California. Initial and boundary meteorological conditions were provided by the North American Regional Reanalysis (NARR) dataset [Mesinger et al., 2006]. All simulation durations were 30 hours including 6 hours of model spin up. The model also incorporated 3-D analysis nudging every three hours in the 36-km domain.

The specific combination of land surface models (LSMs) and PBL schemes that yielded the best comparison with PBL heights retrieved from the wind profilers [Bianco and Wilczak, 2002; Bianco et al. 2008] in the Central Valley varied with season and location. An evaluation of WRF meteorology was performed using data for the Sacramento (SAC), Chowchilla (CCL), Chico (CCO) and Lost Hills (LHS) (sites are shown in Figure 10). For summer and early fall, the combination of the five-layer thermal diffusion LSM (5-L LSM hereafter) and the Mellor-Yamada-Janjic (MYJ) PBL scheme [Mellor and Yamada, 1982; Janjić, 1990] performed best. For example, for the summer month of June 2010 (due to profiler data availability, 2010 data are used for some sites and months), the 5-L LSM and MYJ combination (root mean square (RMS) errors = 280 - 290 m, also shown in Figure 11) performed better than the Noah LSM and MYJ combination (RMS errors = 400 - 450 m) for the SAC and CCL sites. This is likely due to the fact that the 5-L LSM includes irrigation to actively manage soil moisture as a function of land cover types. For winter, the Noah LSM and MYJ combination performed well because the more complicated Noah LSM handles energy balance better when precipitation is the dominant source of moisture. The one exception is that of the LHS site during summer and early fall where the 5-L LSM and Yonsei University (YSU) PBL scheme combination performed better than the 5-L LSM and MYJ combination. For example, for June 2010, the RMS error for the 5-L LSM and MYJ combination (526 m) was significantly larger than that of the 5-L LSM and YSU combination (359 m). We speculate that the 5-L LSM may overestimate soil moisture at the LHS site, reducing PBL height in a manner that is compensated for by the overestimation of PBL height by the YSU scheme. However, we lack the data to test this hypothesis at this time.

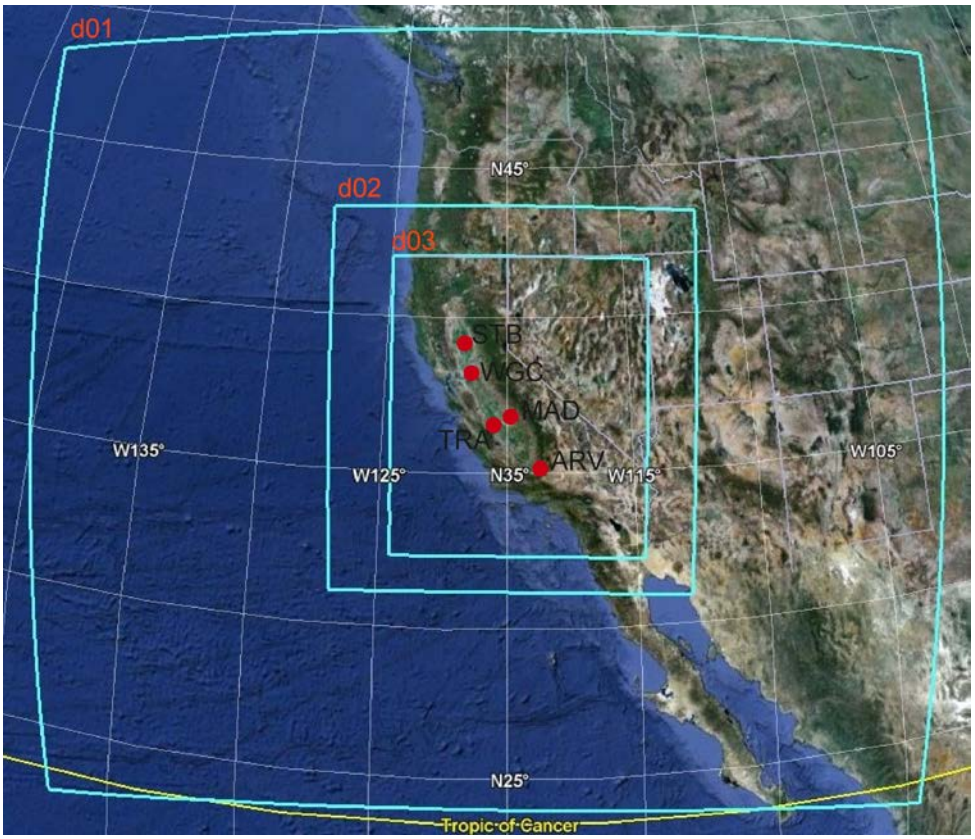

Figure 9. WRF modeling domain configuration with three-level nested domains (d01, d02, and d03 featuring 36, 12, and 4 km resolution, respectively).

2.4. Bayesian Inverse Model

Seasonally varying regional (~10,000 km^2) CH_4 emissions from California's Central Valley are estimated by scaling the high-resolution (~10 km) CH_4 emission maps using a Bayesian inversion model to provide optimal agreement with aggregate mixing ratio data measured at the five-site network. This inversion approach expands on the earlier efforts by Zhao et al. [2009], Jeong et al. [2012a] and Jeong et al. [2012b] and is described here with some background information of Bayesian inversion.

Bayesian inverse modeling for quantifying GHG emissions attempts to gain useful information from measured signals to infer most probable surface emissions based on known *a priori* emissions, which are usually available as bottom-up inventories (e.g., CARB GHG emission inventory). In many cases, inversion efforts are made to estimate optimal parameters such that when those estimated parameters are applied, predicted signals based on *a priori* emissions agree well with measured signals. In this study, the parameters optimized in the inversion are non-dimensional emission factors (or scaling factors) which are multiplied by the *a priori* knowledge (i.e., *a priori* CH_4 emissions) to best match atmospheric CH_4 measurements. Therefore, the measured signal at the receptor can be related to the unknown scaling factor as

$$\textit{measured signal} \cong \textit{predicted signal} \times \textit{scaling factor}$$

where $\cong$ denotes relation, and the predicted signal is a product of footprints and *a priori* emissions and needs to be adjusted by scaling factors to be consistent with measured signals.

This relationship can be formulated as in Gerbig et al. [2003], Lin et al. [2003], Zhao et al. [2009], and Jeong et al. [2012a] and be written as

$$\mathbf{c} = \mathbf{K}\boldsymbol{\lambda} + \mathbf{v}, \qquad (1)$$

where **c, K**, and $\boldsymbol{\lambda}$ represent the measured signal, predicted signal, and scaling factor, respectively. More precisely, **c** is an $n \times 1$ measurement vector (n is the number of measurements; e.g., 3-hourly mixing ratios), $\mathbf{K} = \mathbf{FE}$ (an $n \times k$ matrix), $\boldsymbol{\lambda}$ is a $k \times 1$ state vector for scaling factors (k is the number of regions or sources), which is used to adjust emissions from sources or regions, and **v** is a vector representing the model-data mismatch with a covariance matrix **R**. We model **R** as a diagonal matrix to represent the total variance associated with all error sources such as the measurement error and the transport error. We describe the model-data mismatch matrix in detail in Section 2.5.

As described previously, the footprint **F** is an $n \times m$ matrix, where m is the number of grid cells (~ 10 km × 10 km) in the emission maps. In this linear equation, **F** links the surface emissions **E** to the signal **c** at the measurement site. In other words, the footprint transforms the surface emissions to a mixing ratio enhancement observed at the measurement site. Each column ($n \times 1$ vector) of **F** (total number of columns = m) describes the contribution of surface emissions in a given grid cell (10 km × 10 km) to the signal at the measurement site during the entire prediction period. The emission map **E** is an $m \times k$ matrix where each column ($m \times 1$ vector) has non-zero values only for the grid cells belonging to the kth region or source so that the kth column represents the emissions from the kth region or source. Therefore, **K**, product of **F** and **E,** represents the predicted signals. The kth column of **K** contains the contribution of the kth region or source to the total signal at the receptor for the entire prediction period (e.g., each month). Now, the product of **K** and $\boldsymbol{\lambda}$ yields an $n \times 1$ vector of posterior (optimized) signals, which is the same size as **c**.

Following the Gaussian assumptions, the posterior (optimized) estimate for $\boldsymbol{\lambda}$ is

$$\boldsymbol{\lambda}_{post} = \left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{K} + \mathbf{Q}_\lambda^{-1}\right)^{-1}\left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{c} + \mathbf{Q}_\lambda^{-1}\boldsymbol{\lambda}_{prior}\right) \qquad (2)$$

where $\boldsymbol{\lambda}_{\text{prior}}$ is the *a priori* estimate for $\boldsymbol{\lambda}$ (initially set to one for all elements), and $\mathbf{Q}_\lambda$ is the error covariance matrix ($k \times k$) associated with $\boldsymbol{\lambda}_{\text{prior}}$. The corresponding posterior covariance for $\boldsymbol{\lambda}$ is $\mathbf{V}_{post} = \left(\mathbf{K}^T\mathbf{R}^{-1}\mathbf{K} + \mathbf{Q}_\lambda^{-1}\right)^{-1}$.

Once the posterior (optimized) scaling factor $\boldsymbol{\lambda}_{post}$ is solved using Equation (2), $\boldsymbol{\lambda}_{post}$ is multiplied by *a priori* emissions to yield posterior emissions (i.e., adjusted emissions), which we use to evaluate the current State emission inventory (e.g., CARB inventory). For example, when $\boldsymbol{\lambda}_{post}$ (13 × 1 vector or posterior scaling factors for 13 regions in Figure 5) from Bayesian region analysis is multiplied by the corresponding annual emissions (e.g., Table 2) for the 13 regions, we obtain the posterior emission for each region. Statistically, posterior emissions are the most probable emission estimates given the available data, prior knowledge, and assumptions we made. Summing the posterior emissions over all 13 regions yields posterior State total emissions. Similarly, posterior uncertainty can be obtained by multiplying the annual *a priori* emissions by the square root of the diagonal elements of $\mathbf{V}_{post}$.

To determine optimal emissions, we use the inversion method at a monthly temporal scale based on the two CH_4 *a priori* emission models. Because the measurements sites are located in California's Central Valley that includes such uncertain CH_4 emission sources as rice agriculture, livestock and natural gas fields, we use 50% uncertainty in our *a priori* emission models [Jeong et al., 2012a]. The inverse modeling approach is applied in two phases as in Bergamaschi et al. [2005] and Jeong et al. [2012a]. After a first inversion, the second (final) inversion uses data points that are accepted by applying the selection criteria $|\mathbf{c}_i - (\mathbf{K}\boldsymbol{\lambda})_i|^2 < \alpha\mathbf{R}_i$, where α is a fixed value. In other words, after the first inversion the squared difference between measurements and scaled predicted signals should be less than a threshold (αR_i). In this study we use 2 or 3 for α depending on the season although the choice of α does not make a significant difference in emission estimation. Described previously, **K** ($n \times k$ matrix) represents the predicted signals for k number of regions or sectors, and thus the product of **K** and $\boldsymbol{\lambda}$ yields an $n \times 1$ vector of scaled or optimized predicted signals for n number of data points (n 3-hourly data in this study). The kth column of **K** contains the contribution of the kth region or source to the total signal at the receptor for the entire prediction period (e.g., each month). As in the first inversion, the final inversion is performed using the original *a priori* emission maps, and therefore the first inversion is used as a data selection tool for the atmospheric observations.

2.5. Uncertainty Analysis

The uncertainty in the model-measurement differences controls the relative weighting of the prior flux estimates and the measured data in the inversion, adjusting posterior CH_4 emissions relative to *a priori* emissions. Following Gerbig et al. [2003], Zhao et al. [2009], Göckede et al. [2010], and Jeong et al. [2012a], the model-measurement mismatch matrix, **R** (an $n \times n$ matrix), is represented as the linear sum of uncertainties from several sources and modeled as a diagonal matrix:

$$\mathbf{R}_i = \mathbf{S}_{part} + \mathbf{S}_{aggr} + \mathbf{S}_{bkgd} + \mathbf{S}_{transPBL} + \mathbf{S}_{transWIND},$$

where the particle number error (S_{part}) is due to the finite number of released particles at the receptor location while the aggregation error (S_{aggr}) arises from aggregating heterogeneous fluxes

within a grid cell into a single average flux. The background error (S_{bkgd}) is due to the uncertainty in estimating the background contribution to the CH_4 measurements at the receptor. $\mathbf{S}_{transWIND}$ and $\mathbf{S}_{transPBL}$ represent the uncertainty in CH_4 mixing ratios caused by the errors in wind speeds and directions, and the errors in PBL heights, respectively. For the aggregation error ($\mathbf{S}_{aggr}$), we adopt the result from Jeong et al. [2012a] and use 11% of the background-subtracted mean signal. The background error ($\mathbf{S}_{bkgd}$) is estimated by combining (in quadrature) the RMS error in the estimation of the 3-D curtain (similar to that used in Jeong et al., 2012b) and the standard error of 500 WRF-STILT background samples, which were calculated as an average for each month during September 2010 – June 2011. Recall that we release 500 particles backwards in time for 7 days, and each particle is associated with a background value at its final location. The background errors were estimated to be 17 – 25 ppb depending on the season and measurement site. Only time points for which more than 80% of the particles reached the western boundary of the domain (130°W) were included in the study.

To estimate the uncertainty in predicted CH_4 signals due to errors from modeled PBL heights ($\mathbf{S}_{transPBL}$) and winds ($\mathbf{S}_{transWIND}$), we evaluated WRF model errors in winds and PBL heights and then calculated the RMS difference in CH_4 signals obtained from simulations with and without input of an additional stochastic component of wind and PBL errors in STILT. As described previously, we evaluated PBL heights (*Zi*) and winds at four profiler sites (Figure 10): CCO, SAC, CCL, and LHS. Wind and *Zi* measurements from the closest profiler to the GHG measurement site were used to evaluate WRF simulations. For example, most relevant to the ARV GHG measurement site, we compared *Zi* from WRF with measurements from the LHS profiler. For the MAD and TRA GHG sites, we used wind profiler data from the CCL site. As in Jeong et al. [2012a, 2012b], we assume that the RMS scatter in predicted versus measured *Zi* can be represented as the sum of squares of measurement uncertainty [~ 200 m, Dye et al., 1995] and WRF model uncertainty. For comparison between WRF and profiler measurements, we used data for May 2010, June 2010, October 2010, and January 2011 to represent spring, summer, fall, and winter seasons, respectively. Due to data availability we used 2010 data for spring and summer except for the CCO site for which May and June 2011 data were used. For the LHS site, we used September 2010 data for fall because the LHS profiler data were not available after September 2010. Thus, we used the result from the CCL site for the LHS site after September 2010. The WRF simulated *Zi* was generally consistent with the measured *Zi*. As an example, Figure 11 shows the *Zi* comparison result between profiler measurements and WRF predictions for the month of June.

Following Jeong et al. [2012a, 2012b], we computed CH_4 signals (C_{CH4}) based on the perturbation in *Zi* (20%) to estimate the sensitivity of C_{CH4} to *Zi* (i.e., dC_{CH4}/dZi) as a first order approximation. Then we applied the inferred RMS errors (in units of meter) in the WRF-STILT model to dC_{CH4}/dZi to estimate errors (in ppb) associated with *Zi* for each season and each site. The estimated uncertainties ranged from ~5 ppb to over 200 ppb depending on the season and site, yielding large errors during winter and small errors during summer. For instance, the ARV and MAD sites with the mean background-subtracted signal of ~500 ppb in January showed large errors associated with *Zi* (~200 ppb). In June, the uncertainties due to *Zi* errors in the ARV and MAD sites were relatively small (56 and 35 ppb, respectively) although the mean signals were also low (125 and 105 ppb).

Uncertainty in modeled CH_4 signals due to errors in modeled winds was estimated by comparing WRF-simulated winds and measured winds from the four wind profiler sites for a total of four selected months as in the case of *Zi* (Figure 10). When we compared WRF-simulated winds with profiler-measured winds at the available levels of profilers near the surface (~200 m above mean sea level), the RMS errors in the wind U and V components varied depending on the season and measurement location. For the SAC profiler site (most relevant to WGC), the RMS errors for the wind U/V components were 3.42 (fit slope = 1.00±0.03) / 2.95 (fit slope = 1.13±0.02), 2.89 (1.39±0.11) / 4.96 (1.41±0.11), 3.37 (1.04±0.04) / 3.11 (1.15±0.02), and 2.87 (0.98±0.03) / 2.88(1.05±0.03) m s^{-1} for October, January, May and June, respectively. For the CCL site (most relevant to MAD and TRA), we used data for October and January only because profiler data were not available for spring and summer 2011. The RMS errors for the U/V components were 3.77 (fit slope = 0.96±0.03) / 3.48 (fit slope = 1.04±0.03) and 2.76 (1.01±0.04) /2.91 (1.32±0.05) m s^{-1} for October and January (later we used the SAC site results for the other months to perform STILT ensemble runs). We evaluated winds at the CCO site for the months of May and June 2011 when CH_4 measurements were made at the STB site near the CCO site. The wind U/V RMS errors were 4.22 (fit slope = 1.03±0.04) / 5.99 (fit slope = 1.14±0.03) and 3.17 (0.95±0.03) / 4.45 (1.06±0.03) m s^{-1} for May and June, respectively. Since profiler wind data for the LHS site were not available after early September 2010, we used results from either SAC or CCL sites to run the STILT model for error quantification. For January when WRF overestimated wind speeds relative to profiler winds, we removed outliers (data points corresponding to > 2σ of measured wind speed) to avoid biases in inverse analyses.

To estimate the effect of uncertainty in CH_4 signals due to winds ($\mathbf{S}_{transWIND}$) and particle number ($\mathbf{S}_{part}$), we ran the STILT model 10 times and computed ensemble signals for a given site and month (a total of four selected months as in the *Zi* case). Based on 10 ensemble runs, we estimated the RMS difference about the mean of the ensemble signals for each model time step and use the monthly average RMS as the combined uncertainty due to wind and particle number errors. Following the method in Jeong et al. [2012a; 2012b], propagating a random wind component of the wind velocity error through STILT yielded a signal variation of 1 – 15 ppb depending on the season and site. As with the *Zi* case, the errors due to winds were higher during winter (8 – 15 ppb) than during summer (~ 2 ppb).

Following Zhao et al. [2009] and Jeong et al. [2012a; 2012b], we assumed that all of the errors are independent. The errors were combined in quadrature to yield a total expected model-data mismatch error. Depending on the month and measurement location, the errors ranged from 20 to 230 ppb, which are approximately 30 – 60% of the background-subtracted mean signal. The total error was particularly large (100 – 230 ppb) during winter in the ARV and MAD sites where the mean signal was also high (220 - 520 ppb).

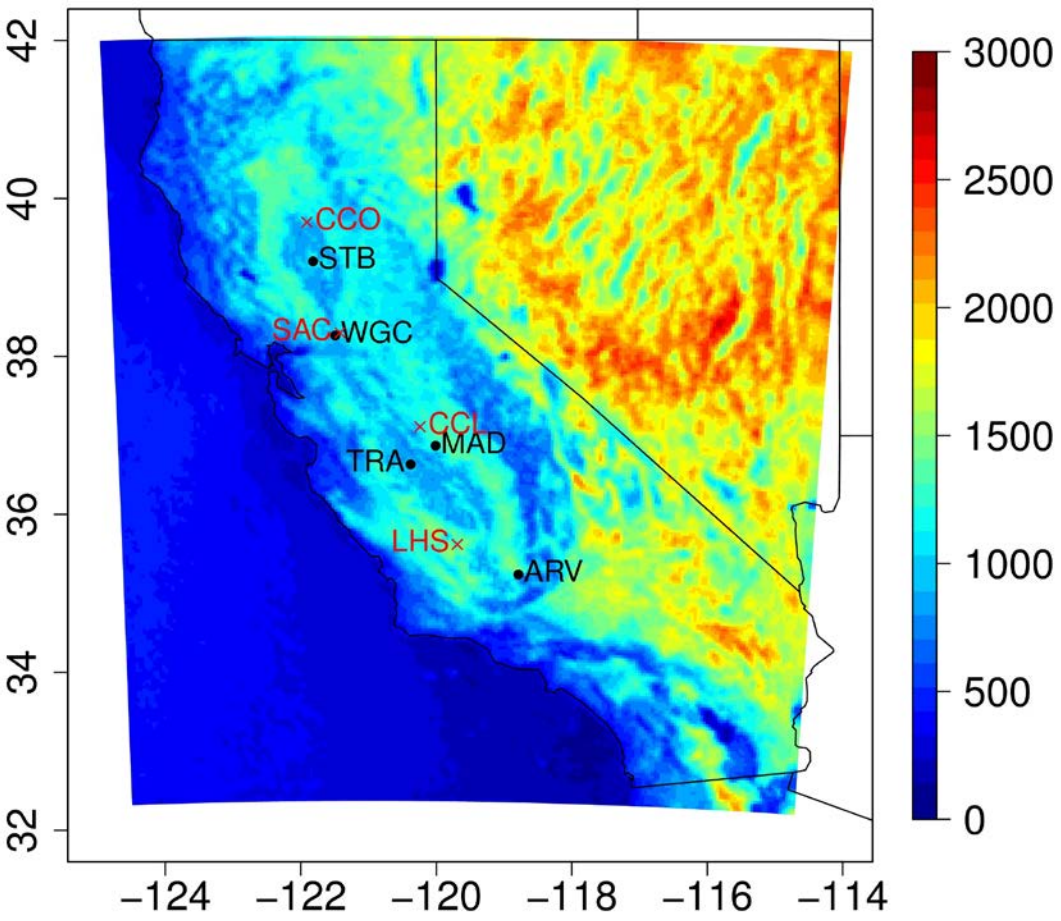

Figure 10. Location of GHG measurement sites (black) and wind profiler sites (red) in the Central Valley with predicted monthly mean PBL heights (m) for June 2011, 14:00 local standard time (LST) shown in color.

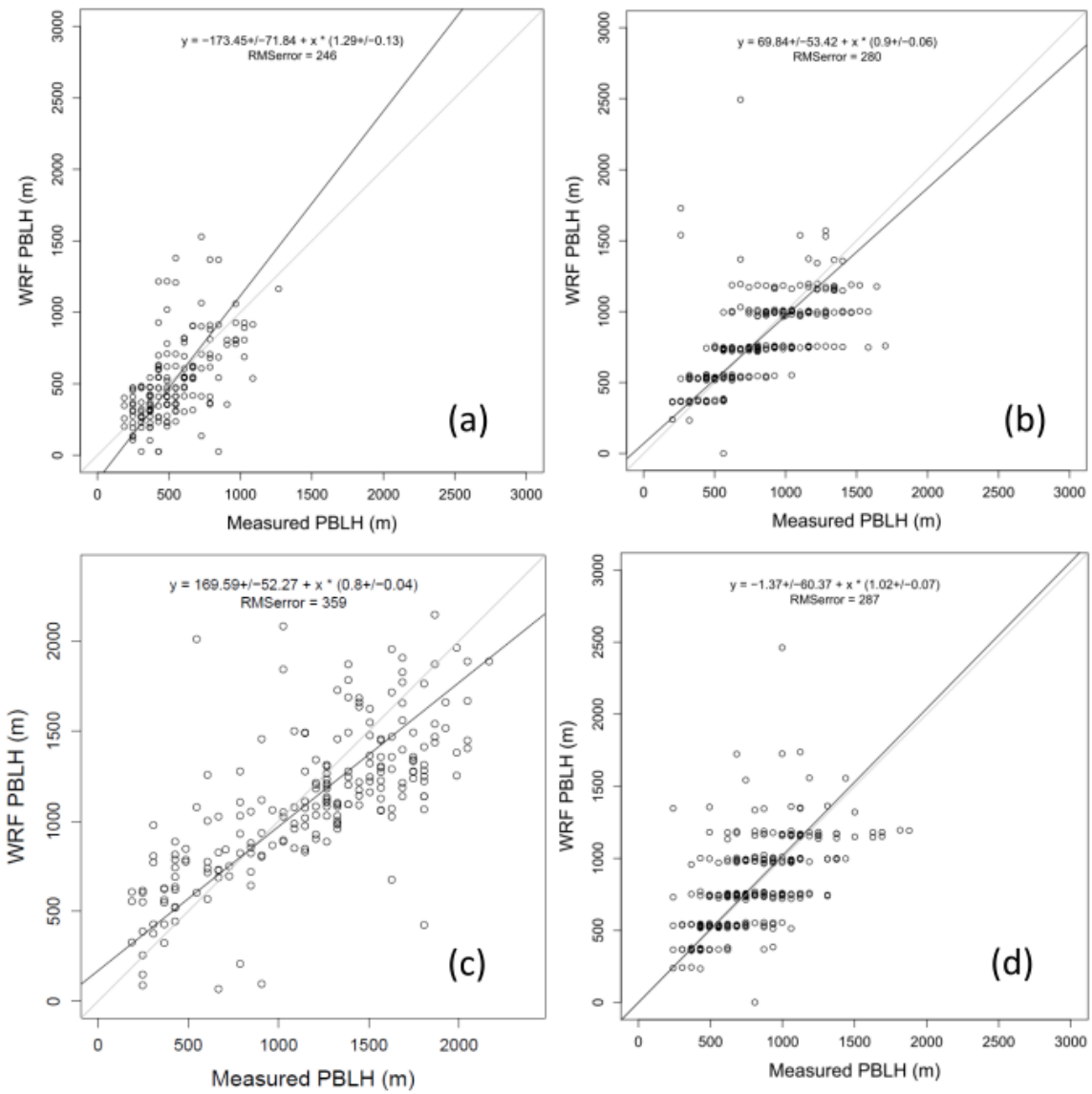

Figure 11. Comparison of measured and predicted *Zi* during the month of June for (a) CCO, (b) CCL, (c) LHS, and (d) SAC. For CCO, data from June 2011 data are used while the other sites use data from June 2010. For this summer month, the 5-L LSM scheme was used for all sites. In terms of the PBL scheme, the MYJ scheme was used for all sites except for LHS where the YSU scheme was used.

3. Results and Discussion

3.1. CH_4 Mixing Ratio

Figure 12 shows the 3-hourly background-subtracted CH_4 mixing ratios and initial predictions using the *a priori* CH_4 emissions for the five network sites. Predicted mixing ratios are shown only for the well-mixed periods which are used for inverse analyses. Overall, the predicted signals at all sites show underestimation of CH_4 compared to the measurements although the

prediction captures the synoptic variation of the measured signals. Also, the daytime (noon – afternoon) comparison between the predicted and measured signals is shown in Figure 13. The result during the daytime suggests that there is a clear seasonal variation in CH_4 signals with high variability in winter. The CH_4 predictions at ARV and MAD are more consistent with the measurements than at the other measurement sites. It is likely due to the fact that the footprints contributing to ARV and MAD cover areas with CH_4 emissions from dairy livestock and oil and gas extractions, for which the California-specific emission model estimate significantly higher emissions compared to EDGAR42. Note that based on the California-specific model, the CH_4 emission sum for Region 12, which includes ARV, is ~1.6 times higher than that of EDGAR42. This suggests that emission estimates for Region 12 from both emission models are low. The comparison result in STB indicates that late spring and summer emissions from rice agriculture are significantly lower than actual emissions. The DNDC model suggests that CH_4 emissions from rice agriculture in Region 6 become strong starting in June with an emission sum of 3.4 Tg CO_2eq yr^{-1} and peaking in August with emissions equating to 4.6 Tg CO_2eq yr^{-1}. For WGC, the predicted signals are significantly lower than the measurements, showing similar results to those shown in Jeong et al. [2012a].

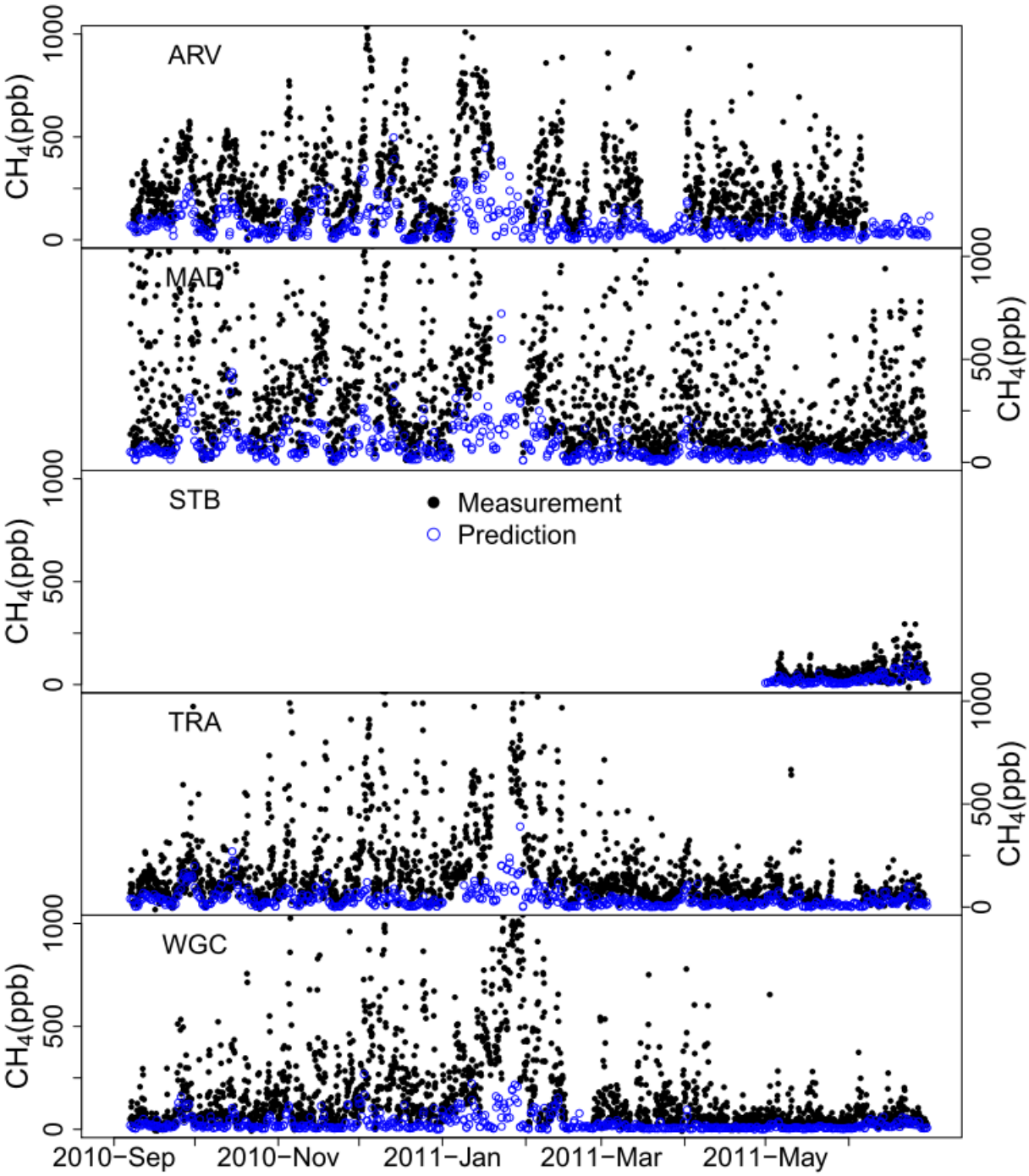

Figure 12. Time series of measured and predicted mixing ratios at the five network sites during September 2010 – June 2011. The measurements are shown for both day and night times while the predictions are shown for only noon-afternoon well-mixed periods. The prediction was made based on the California-specific emission maps.

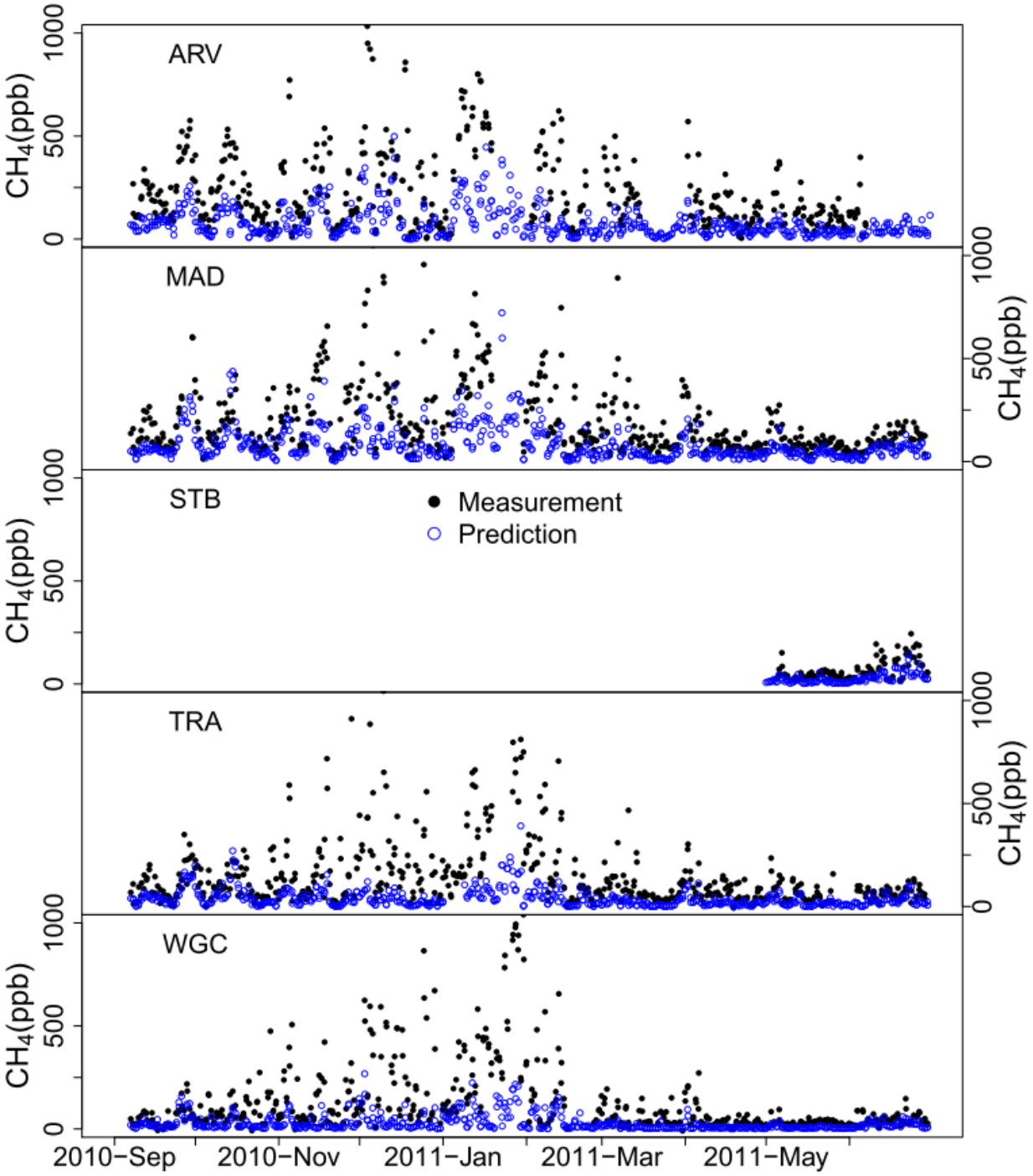

Figure 13. 3-hourly time series of measured and predicted mixing ratios at the five network sites during noon-afternoon well-mixed periods. The prediction was made based on the California-specific emission maps.

3.2. Footprints

When footprints for all five sites are combined, the sensitivity of the measurement sites to surface emissions is significantly improved, as compared to the results with one site only. The combined footprints from multiple sites extend the sensitivity over most of the Central Valley. Figure 14 shows the average CH_4 footprints for May – June 2011 and compares the average footprints between the cases using one site and multiple sites. The averaged footprint in each grid cell shown in Figure 14 is the average of all noon-afternoon footprint values for the grid cell during May – June 2011. Thus, this average footprint map represents the overall behavior of particles released from the measurement sites during the given period. In particular, the averaged footprints from the multiple sites show the combined sensitivity of measurements at the receptors to the surface emissions that contribute to measured signals (see Section 2.3 for details on footprint calculations). The significance is clear in Figure 14 where the averaged footprint from a single tower shows limited sensitivity (mostly strong in the Northern Central Valley) while the footprint from multiple sites shows strong sensitivity in the entire Central Valley. Recall that footprints (in units of concentration / surface emission flux) show how sensitive the signal at the measurement site is to the emission flux at each grid cell (~10 km × 10 km) within the domain of study. Therefore, strong sensitivity over a larger region suggests that the inverse model can constrain emissions for the extended region. One of the significant implications of this study is that the measurement sites in the Central Valley have limited ability to constrain CH_4 emissions in the Southern California region due to weak sensitivity.

Figure 15 shows average footprints for other seasons. There is a clear seasonal pattern for the distribution of footprints, which is important to attribute signals to different emission sources for each season. Overall, the seasonal footprints are strong in the north-south direction in the Central Valley although footprints are strong in the west-east direction near the WGC site for some seasons. During summer, footprints are also strong from the San Francisco Bay Area to the west of the WGC tower due to the dominance of land-ocean winds. Also, there is a shift of footprints toward north-south winds from west-east winds near the WGC site during the transition seasons of spring and fall, which is a similar result to that described in Jeong et al. [2012a] and Jeong et al. [2012b]. Depending on the season, the footprints reach the Southern California Air Basin (i.e., Region 10), which allows for constraining important urban emissions.

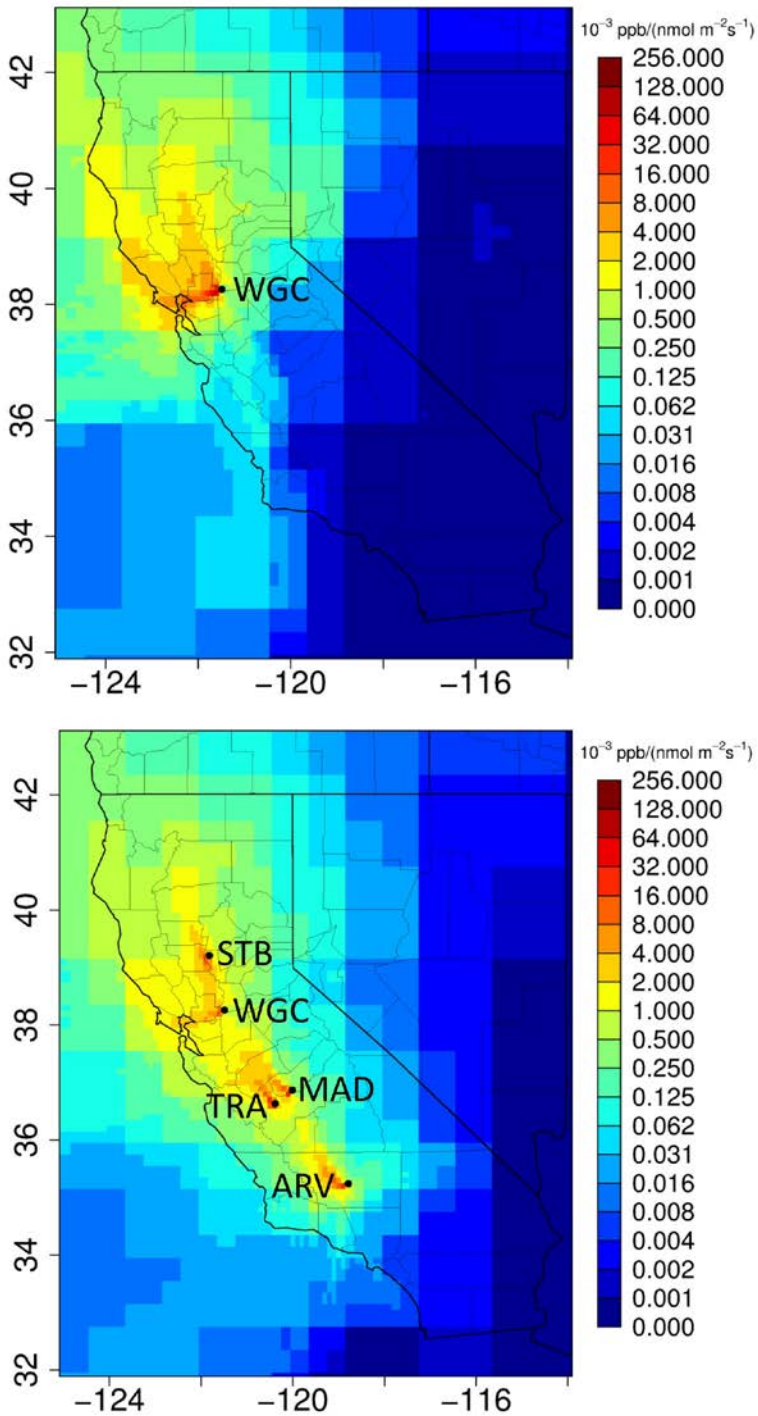

Figure 14. Averaged footprints during the noon-afternoon hours for (a) the WGC site and (b) all five sites during May – June 2011.

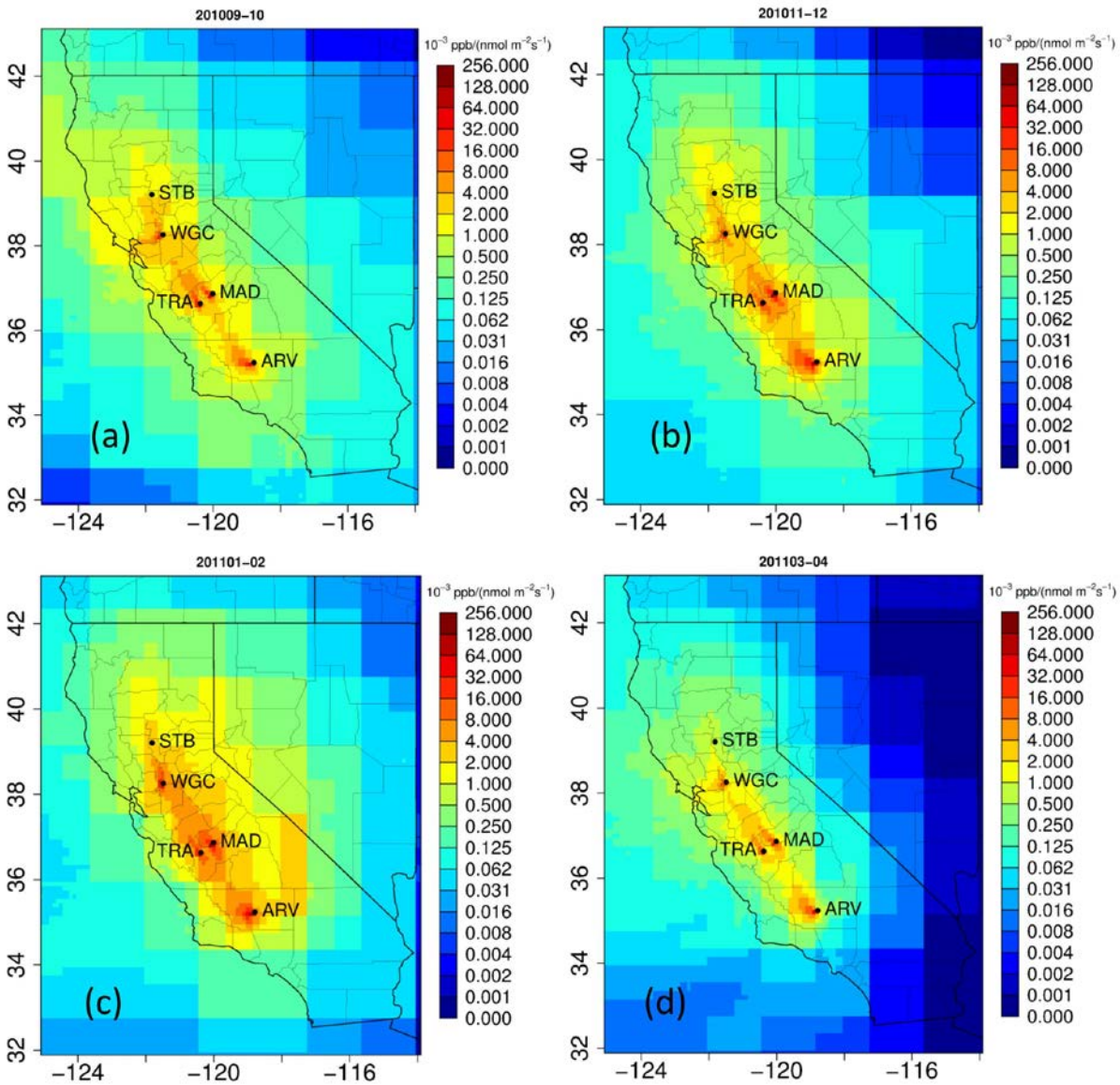

Figure 15. Seasonal mean footprints during the noon-afternoon hours for (a) September – October 2010, (b) November – December 2010, (c) January – February 2011 and (d) March – April 2011.

3.3. Bayesian Inverse Analysis

3.3.1. Linear Analysis

Bayesian inverse analysis was conducted using two independent emission models: California-specific and EDGAR42 emission models. Using each emission model, we performed Bayesian inversion to estimate optimized emissions for both each emission region defined in Figure 5 and each emission source (e.g., livestock). The Bayesian inversion for estimating region emissions is

referred to as Bayesian region analysis (or simply region analysis). Similarly, the latter is called Bayesian source analysis (or source analysis). Note that for both analyses we use a scaling factor Bayesian inversion (SFBI) technique introduced in Section 2.4. Therefore, we obtain four different emission estimates from Bayesian inversions. For the region analysis, we first optimize scaling factors for the 13 regions in Figure 5 using the SFBI technique (Equations 1 – 2) and then multiply *a priori* emissions (emissions are organized by region) by the optimized scaling factors to obtain optimized (posterior) regional emissions. Similarly, the source analysis optimizes scaling factors for emission sources, which are multiplied by *a priori* emissions (emissions are organized by source sector) to obtain posterior source emissions.

Figure 16 shows the comparison result between measured CH_4 signals and those predicted based on the California-specific emission model using a chi-squared (fitexy) linear regression analysis [Press et al., 1992] during May 2011. In the figure, outliers are removed after initial inversion [Bergamaschi et al., 2005].

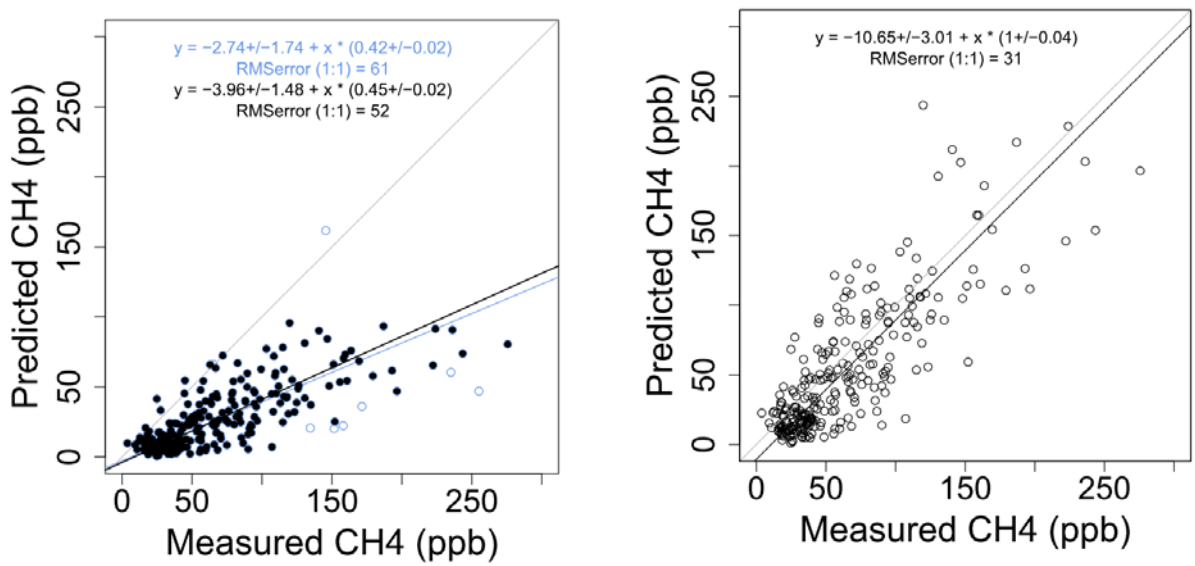

Figure 16. Comparison of California-specific predicted vs. measured CH_4 signals during May 2011 before (left) and after (right) inverse optimization. The light blue circles indicate those removed after the first inversion.

Table 5 summarizes the linear analysis results before and after Bayesian region inversion where emissions for each sub-region are scaled based on the optimized scaling factors for 14 sub-regions including the region outside California. Depending on the month, predicted CH_4 signals using the California-specific emission model are typically 30 – 50% of measurements before inversion, while EDGAR42 prior signals are 20 – 40% of measurements before inverse optimization. After inversion with the 50% error assumption in the prior emissions, the posterior CH_4 signals based on the California-specific emission model are consistent with the measurements for most of the months, while the posterior signals from EDGAR42 are still systematically lower than the measurements. This suggests that EDGAR42 emissions in the Central Valley are systematically too low where footprints are strong.

Table 5. Linear Analysis Results Before and After Bayesian Region Inversion

		Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June
						Before Inversion					
CA[a]	Slope	0.47±0.02	0.41±0.05	0.42±0.03	0.42±0.05	0.40±0.05	0.3±0.03	0.35±0.02	0.43±0.02	0.42±0.02	0.45±0.04
	RMSE[c]	70	108	155	233	303	149	108	78	61	63
ED42[b]	Slope	0.28±0.03	0.30±0.04	0.27±0.02	0.22±0.03	0.24±0.03	0.25±0.03	0.23±0.02	0.28±0.01	0.28±0.02	0.36±0.05
	RMSE	90	113	171	254	334	156	119	90	68	69
						After Inversion					
CA	Slope	0.91±0.04	0.89±0.08	0.93±0.05	0.96±0.08	0.94±0.09	0.80±0.08	0.93±0.05	0.97±0.04	1.00±0.04	0.94±0.05
	RMSE	36	61	83	95	165	92	63	44	31	32
ED42	Slope	0.63±0.04	0.69±0.06	0.84±0.05	0.83±0.08	0.76±0.08	0.65±0.06	0.74±0.03	0.83±0.03	0.86±0.04	0.85±0.07
	RMSE	56	62	83	123	185	93	70	45	34	36

[a]California-specific emission model
[b]EDGAR42 emission model
[c]Root mean square error in units of ppb

3.3.2. Bayesian Region Analysis

The Bayesian region analysis estimates scaling factors for sub-region emissions using the Bayesian inverse technique (see Figure 5 for sub-region classification). In this study, a total of 14 scaling factors including the region outside California are solved for a given inversion period. Although the inversions are performed at the monthly temporal scale, inferred CH_4 emissions are reported by season and as a regional sum for the regions where the total emissions are significant and footprints show sensitivity. Since our data do not cover a full year, we divided the 10 months into five bi-monthly seasons.

Figure 17 shows the posterior emission estimates using the California-specific emission model by region and season. Here, posterior emissions represent those emissions that we obtain after multiplying the *a priori* emissions by the optimized (posterior) scaling factors. Statistically, posterior emissions are the most probable emission estimates given the available data, prior knowledge, and assumptions we made. Overall, the inversion results suggest that actual CH_4 emissions are higher than the prior emissions for most of the regions and seasons. In particular, the posterior emissions are significantly higher than the prior in Regions 6, 8, and 12 where the emissions are well constrained. For Region 10 (Southern California region), the posterior uncertainties are slightly reduced for some seasons, suggesting that the measurements in the Central Valley weakly constrain the emissions in Region 10. In Region 7 (Bay Area and surrounding urban area), summer emissions are slightly higher than the other seasons. The significantly higher posterior emissions in Regions 8 and 12 suggest that emissions from livestock sources are significantly higher than the prior. Note that livestock emissions from the California-specific emission model account for 87% and 84% of the total emissions in Regions 8 and 12, respectively. The posterior results in Figure 17 also show that there is a clear seasonal variation in CH_4 emissions. For example, in Region 6 where high emissions are expected from

rice agriculture, the posterior emissions are high during the fall (i.e., fall 2010) and later spring-early summer season. We discuss rice emissions in detail in the following section.

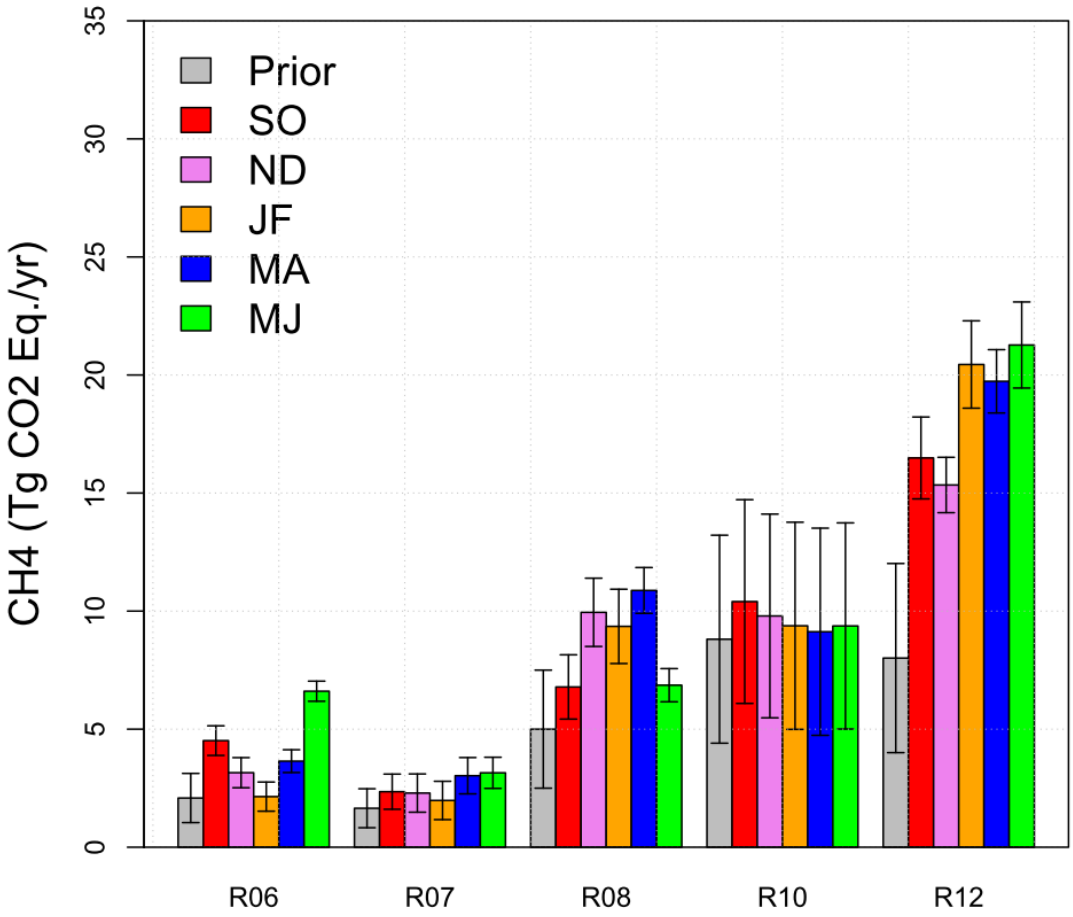

Figure 17. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) by region and season based on the California-specific emission model. Only regions with significant emissions are shown. The annual mean prior (gray bar) represents the annual average of seasonally varying emissions and is compared with posterior seasonal emissions (color bars). SO, ND, JF, MA, and MJ denote September-October, November-December, January-February, March-April, and May-June seasons, respectively.

Figure 18 shows the region analysis results using the EDGAR42 emission model. The inversion results based on EDGAR42 show different emission estimates, yielding generally higher emissions for most of the regions compared to those of the California-specific case. In Region 12, for instance, the EDGAR42-based inverse modeling estimates consistently higher CH_4 emissions than those from the California-specific modeling, although the EDGAR42 prior emission is lower than that of the California-specific model. As shown in Figure 18, the annual average scaling factor for Region 12 is ~3, which yields 15.64±1.47 Tg CO_2eq yr^{-1}.

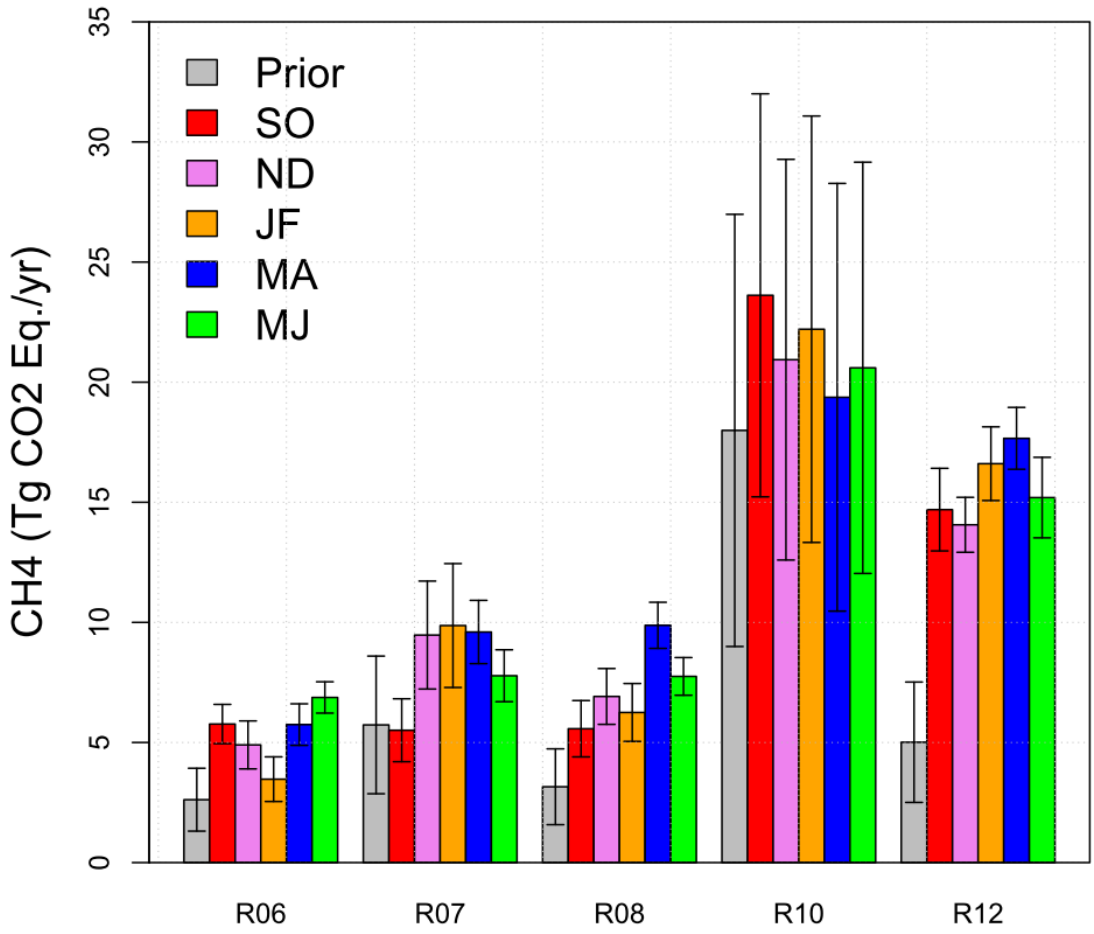

Figure 18. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) by region and season based on the EDGAR42 emission model. Only regions with significant emissions are shown.

Results of the region analysis show that the sum of the posterior emissions for the sub-regions in the Central Valley (i.e., Regions 6, 8 and 12) are similar for the California specific and EDGAR42 *a priori* emission models (Table 6). The current measurement network constrains annual average CH_4 emissions for Regions 6, 8 and 12 to be between 31.43±2.07 and 28.27±2.00 Tg CO_2eq for the California-specific and EDGAR42 emission models respectively, assuming uncorrelated errors between regions. However, there are significant differences in the predominantly urban regions (7 and 10) where the EDGAR42 model shows higher CH_4 emissions than those estimated with the California-specific model: 29.79±8.78 vs. 12.18±4.42. This is because the prior emissions show the same pattern, with EDGAR42 higher than California-specific by a factor of 2.3 (23.7 vs. 10.5 Tg CO_2eq yr^{-1}), and our measurement sites in the Central Valley have relatively weak sensitivity to the urban regions (e.g., Southern California region). This result is consistent with that reported in Jeong et al. [2012a] where the estimated annual emission for Region 7 using EDGAR32 (version 3.2) is higher than that estimated with the California-specific model by a factor of ~4. Although the results using multiple emission models help to characterize the uncertainty associated with estimating emissions at the sub-regional scale, this result demonstrates that additional measurements are required in the San

Francisco Bay and Southern California areas in order to strongly constrain emissions from those urban regions.

Table 6. Comparison of Annual Posterior CH_4 Emissions (Tg CO_2eq) between the EDGAR42 and California-specific Emission Models Based on Bayesian Region Analysis

Emission Model	Region	R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	Total
CA-specific	Emissions	0.1	0.2	0.4	0.3	0.4	4.0	2.6	8.8	0.7	9.6	0.3	18.7	0.2	46.1
	Uncertainty[a]	0.0	0.1	0.2	0.1	0.2	0.6	0.8	1.2	0.3	4.4	0.1	1.6	0.1	4.9
EDGAR	Emissions	0.1	0.4	0.7	0.4	0.5	5.4	8.4	7.3	1.2	21.3	0.4	15.6	0.5	62.2
	Uncertainty	0.0	0.2	0.3	0.2	0.2	0.9	1.7	1.1	0.6	8.6	0.2	1.5	0.2	9.1

[a]Posterior uncertainty = 1σ

3.3.3. Bayesian Source Analysis

We also estimate CH_4 emissions by inferring state-wide scaling factors for each emission source instead of each sub-region. The results show that while posterior emissions from livestock are similar for the California specific and EDGAR42 models, emissions from landfill (solid waste) and natural gas production and use are both substantially higher when using the EDGAR42 prior. Figure 19 shows the source analysis results using the California-specific emission model, and annual average emissions for the State are summarized by emission sector in Table 7. These results are consistent with those of the counterpart inverse analysis for regional emissions. For example, the source inversion suggests that actual emissions from livestock are much higher than the prior. This result agrees with the higher posterior emissions in Region 8 where livestock emissions are dominant (~90% of annual CH_4 emissions).

The source analysis results also indicate that the posterior emissions for crop agriculture are higher during early fall and summer season than the prior, which are consistent with the higher emissions in Region 6, where high biogenic emissions are expected from rice agriculture. This result is similar to that of a recent study based on aircraft CH_4 measurements during the California Research at the Nexus of Air Quality and Climate Change (CalNex) period in summer 2010 [Peischl et al., 2012]. Peischl et al. [2012] estimated annual CH_4 emissions from rice cultivation to be 1.64 – 1.95 Tg CO_2eq. This estimate is based on the rice emission study in a commercial rice field by McMillan et al. [2007] where they estimated annual CH_4 emissions of 26.1 – 31.0 g CH_4-C m^{-2} during October 2001 – October 2002. This estimate is 3.0 – 3.6 times larger than the CARB 2008 inventory (0.54 Tg CO_2eq yr^{-1}) for rice CH_4 emissions. Assuming posterior emissions for July and August (not available in our study) are proportional to the prior and scaling (available) June posterior emissions according to the prior ratios of July and August to June (3.26 Tg CO_2eq / 3.59 Tg CO_2eq and 5.10 / 3.59, respectively), we find that the annual rice emission total is 1.40±0.14 Tg CO_2eq, which is very similar to that of Peischl et al. [2012]. This result suggests that our inverse analysis constrains seasonally-varying rice emissions and demonstrates the capability of the inversion. The slight difference between the estimate by Peischl et al. [2012] and our estimate is possibly due to the difference in emissions during late fall and winter. CH_4 emissions during late fall and winter from McMillan et al. [2007], on which

the estimate by Peischl et al. [2012] is based, are not negligible while our *a priori* rice emissions based on the DNDC model described by Salas et al. [2006] are insignificant and often negative.

While Figure 19 shows that CH_4 emissions from natural gas sources are generally higher than the prior, more measurements are required to effectively constrain natural gas emissions from the large urban areas including the Southern California region.

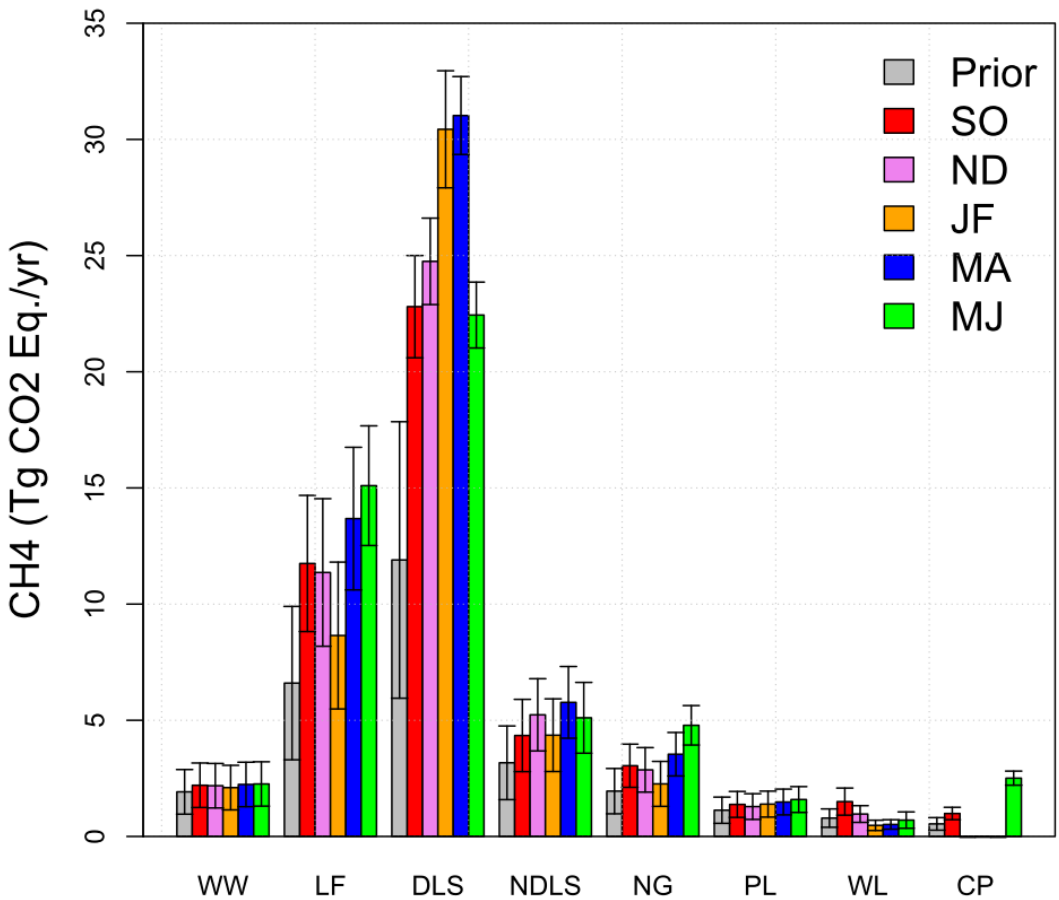

Figure 19. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) for California by source and season based on the California-specific emission model. WW, LF, DLS, NDLS, NG, PL, WL and CP represent wastewater, landfill, dairy livestock, non-dairy livestock, natural gas, petroleum, wetland, and crop agriculture sources, respectively.

Table 7. Annual Posterior CH_4 Emissions (Tg CO_2eq) for California by Source Based on the California-specific Emission Model

Sectors	WW	LF	DLS	NDLS	NG	PL	WL	CP	Total
Emissions	2.20	12.11	26.29	4.96	3.30	1.43	0.83	0.68	51.80
Uncertainty[a]	0.96	3.01	2.30	1.55	0.93	0.56	0.35	0.12	4.44

[a]Posterior uncertainty = 1σ

The source analysis results based on EDGAR42 are shown in Figure 20, where eight major sources (~95% of total emissions) out of a total of 16 sources are compared. The corresponding annual average emissions for the State are summarized by emission sector in Table 8. As one might expect from the region analysis result, the source analysis based on EDGAR42 shows significantly different posterior emissions for some of the sources, compared to the California-specific case. In particular, the annual CH_4 emission estimate for solid waste (equivalent to landfill of the California-specific model) from EDGAR42 is 24.14±4.99 Tg CO_2eq, which is significantly higher than that (12.11±3.01) estimated using the California-specific model. This discrepancy is likely due to the fact that ~70% of landfill emissions are concentrated in the urban regions (Regions 7 and 10), and these urban regions are only weakly constrained by the measurements. This suggests that more measurements are needed to constrain landfill emissions in urban regions.

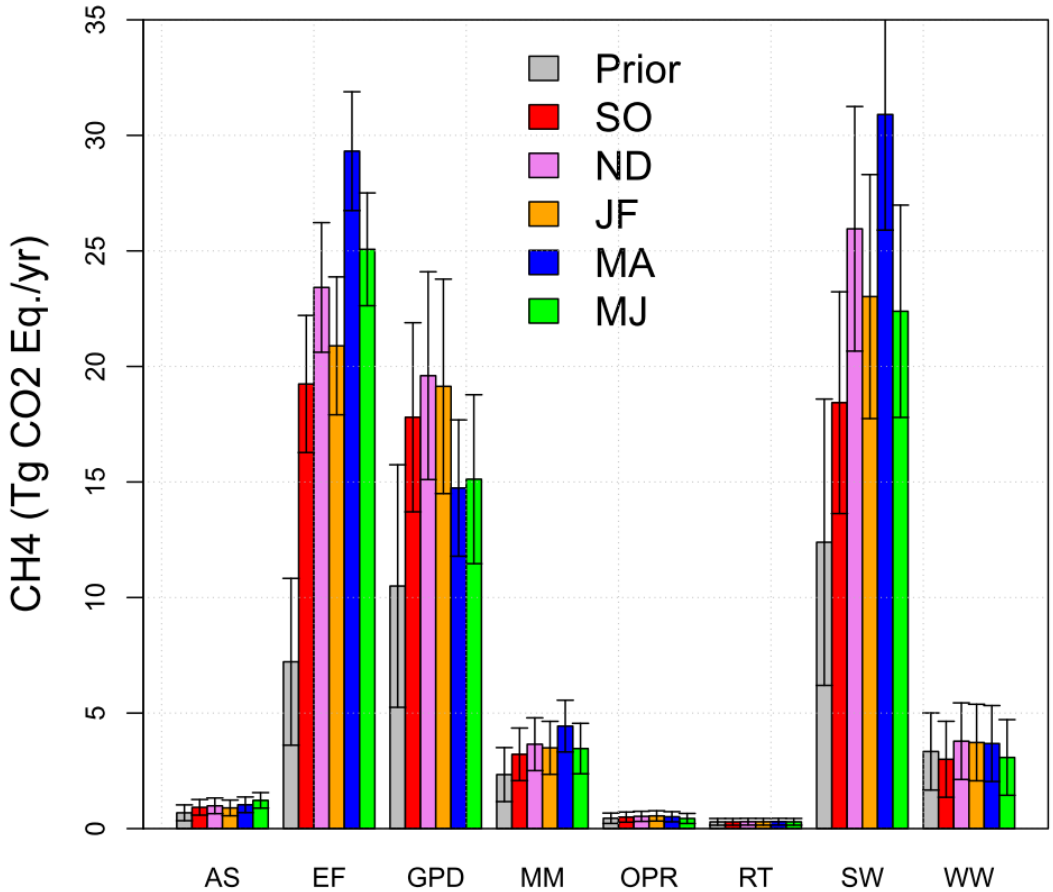

Figure 20. Estimates of posterior CH_4 emissions (Tg CO_2eq yr^{-1}) for California by source and season based on the EDGAR42 emission model. AS, EF, GPD, MM, OPR, RT, SW, and WW represent agricultural soils, enteric fermentation, gas production and distribution, manure management, oil production and refineries, road transportation, solid waste, and wastewater, respectively.

Table 8. Annual Posterior CH_4 Emissions (Tg CO_2eq) for California by Source Based on the EDGAR42 Emission Model

Sectors	AS	EF	GPD	MM	OPR	RT	RT	SW	WW	Others	Total
Emissions	1.01	23.59	17.28	3.65	0.51	0.57	0.30	24.14	3.45	0.67	75.17
Uncertainty[a]	0.34	2.75	3.97	1.13	0.22	0.27	0.15	4.99	1.65	0.61	7.27

[a]Posterior uncertainty = 1σ

Table 9 summarizes the estimated annual CH_4 emissions using the two independent CH_4 emission models. Overall, the EDGAR42-based emission estimates are higher than those of the California-specific model. Also, the source-based inversion analysis estimates slightly higher emissions compared to those of the region analysis. The Bayesian region analysis suggests that the total posterior CH_4 emission for the entire California is 1.44±0.15 times and 1.94±0.28 times higher than the current CARB inventory (32 Tg CO_2eq; CARB, 2011) using the California-specific and EDGAR42 emission models, respectively. The source analysis suggests slightly higher emissions, which are 1.62±0.14 times and 2.35±0.23 times the current CARB inventory for the California-specific model and the EDGAR42 model, respectively.

In addition to the inverse analysis based on measurements in the Central Valley, we used CH_4 emission estimates from Wennberg et al. [2012] where the urban emissions in the Southern California region are better constrained. Our State total CH_4 emission estimates from different inverse analyses are compared with an estimate based on Wennberg et al. [2012]. For California's South Coast Air Basin (the larger Los Angeles metropolitan region, hereafter LA megacity), Wennberg et al. [2012] estimated a total CH_4 emission of 0.44±0.15 Tg yr^{-1}, which is 0.91 – 1.84 times higher than the California-specific prior emission estimate, and 0.4 – 0.9 times the EDGAR42 estimate for the LA megacity. This suggests that the EDGAR42 prior significantly overestimates CH_4 emissions in the LA megacity while the California-specific prior is more likely to underestimate CH4 emission in the region. This also suggests that scaling EDGAR42 emissions by a factor of much greater than 1 based on the inversion results for the Central Valley will substantially overestimate State total CH_4 emissions. Note that annual posterior emissions for the Central Valley from EDGAR42 region analysis are 2.62±0.18 times the EDGAR prior. A more reasonable approach for State total emissions can be estimated by using the posterior emissions from the California-specific model for the Central Valley and other non-urban regions and scaling the California-specific prior emissions for the large urban regions (7 and 10) by a factor of 0.91 - 1.84. This approach estimates total CH_4 emissions of 41.43 – 55.50 Tg CO_2eq yr^{-1}, where the uncertainty is dominated by the urban regions. This result suggests that when this estimate is used, California total emissions are better constrained (factors of 1.30 – 1.74) compared to that from our original region analysis (1.29 – 2.22). This estimate for State total CH_4 emissions is consistent with the results of our inverse analyses (Table 9) except for the EDGAR42 source analysis. Therefore, we do not include the emissions estimated from the source analysis using EDGAR42 in the executive summary because of the significant uncertainty associated with the large urban areas.

Table 9. Summary of Estimated Annual CH_4 Emissions (Tg CO_2eq; 100-year GWP = 21 g CO_2eq / g CH_4) for California

Emission Model	Region Analysis	Source Analysis
California-specific	46.10±4.93	51.80±4.44
EDGAR42	62.18±9.07	75.17±7.27

4. Conclusions and Recommendations

- The current GHG network constrains annual average CH_4 emissions from California's Central Valley to be between 31.43±2.07 Tg CO_2eq and 28.27±2.00 Tg CO_2eq for the California-specific and EDGAR42 emission models respectively, showing consistency between the two independent models. Similarly, emissions for the entire State from livestock (which are predominantly located in the Central Valley) are estimated to be 31.25±2.77 Tg CO_2eq and 27.24±3.13 Tg CO_2eq from the California-specific and EDGAR42 emission models, respectively.

- While significant error reductions are obtained in California's Central Valley, emissions from other regions remain uncertain, with the ratio of emissions to the current California CH_4 emission inventory (32 Tg CO_2eq yr^{-1}) ranging from 1.44±0.15 and 1.94±0.28 from the Bayesian region analysis based on the California-specific and EDGAR42 emission models. Additional tower measurements in the San Francisco Bay and Southern California coastal areas are expected to constrain those emissions.

- Noting the large uncertainty in urban emissions estimated from measurements in the Central Valley, emissions from large urban areas (San Francisco Bay Area and Southern California region) are estimated based on a recent study [Wennberg et al., 2012] in the larger Los Angeles metropolitan region to better constrain large urban emissions. Combined with the emissions from the Central Valley and other non-urban regions, State total CH_4 emissions are estimated to be 1.30 – 1.74 times larger than the current State total CH_4 emissions (32 Tg CO_2eq yr^{-1}) where the uncertainty is dominated by uncertainty in the urban regions. This further suggests that additional measurements in the San Francisco Bay and Southern California areas are required to constrain those emissions.

- Data from the current CH_4 measurement network are effective for use in constraining emissions from different regions of California's Central Valley but cannot be used to uniquely attribute emissions to specific source sectors. Additional measurements of source-specific tracers (e.g., CO, VOCs, and potentially CH_4 isotopes) will help separate different sources of CH_4.

- Currently, uncertainty in the inverse model estimates of CH_4 emissions for regions containing measurement sites are dominated by uncertainty in the meteorological

modeling of trace gas transport (in winter) and estimation of background signals (in summer). Additional work is needed to identify the source of these errors and reduce them.

5. References

Bergamaschi, P., M. Krol, F. Dentener, A. Vermeulen, F. Meinhardt, R. Graul, M. Ramonet, W. Peters, and E. J. Dlugokencky (2005), Inverse modelling of national and European CH_4 emissions using the atmospheric zoom model TM5. Atmos. Chem. Phys., 5, 2431–2460.

Bianco, L., and J. Wilczak (2002), Convective boundary-layer depth: improved measurement by Doppler radar wind profile using fuzzy logic methods. J. Atmos. Oceanic Tech., 19, 1745-1758.

Bianco, L., J. Wilczak, and A. White (2008), Convective boundary layer depth estimation from wind profilers: statistical comparison between an automated algorithm and expert estimations. J. of Atmos. and Oceanic Tech., 25, 1397–1413, DOI: 10.1175/2008JTECHA981.1

CARB (2010), California Greenhouse Gas Emissions Inventory. California Air Resources Board Staff Report. (http://www.arb.ca.gov/cc/inventory/inventory.htm)
CARB (2011), California Greenhouse Gas Emissions Inventory. California Air Resources Board Staff Report. (http://www.arb.ca.gov/cc/inventory/inventory.htm)

Dye, T.S., C. G. Lindsey, and J. A. Anderson (1995), Estimates of mixing depth from "boundary layer" radar profilers. Preprints from the 9th Symposium on Meteorological Observations and Instrumentation, Charlotte, NC, March 27-31, 156-160 (STI-94212-1451).

EDGAR42 database: Emission Database for Global Atmospheric Research (EDGAR), release version 4.2, http://edgar.jrc.ec. europa.eu, European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL), 2012.

Gerbig, C, J. Lin, S. Wofsy, B. Daube, A. E. Andrews, B. Stephens, P. S. Bakwin and C. Grainger (2003), Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. J. Geophys. Res., 108(D24), doi:10.1029/2003JD003770.

Gimson, N. R. and Uliasz, M. (2003), The determination of agricultural methane emissions in New Zealand using receptor-oriented modeling techniques. Atmos. Environ., 37, 3903-3912.

Göckede, M., A. M. Michalak, D. Vickers, D. P. Turner, and B. E. Law (2010), Atmospheric inverse modeling to constrain regional-scale CO_2 budgets at high spatial and temporal resolution. J. Geophys. Res., 115, D15113, doi:10.1029/2009JD012257.

Hofman, D. J., J. H. Butler, E. J. Dlugokencky, J. W. Elkins, K. Masarie, S. A. Montzka, and P. Tans (2006), The role of carbon dioxide in climate forcing from 1979 - 2004: Introduction of the Annual Greenhouse Gas Index. Tellus B, 58B, 614-619.

Houweling, S., T. Kaminski, F. Dentener, J. Lelieveld, and M. Heimann (1999), Inverse modeling of methane sources and sinks using the adjoint of a global transport model. J. Geophys. Res., 104(D21), 26137-26160.

IPCC (1995), IPCC Second Assessment Report (SAR), “Climate Change 1995”.

Janjic, Z. I. (1990), The step-mountain coordinate: physical package. Mon. Weather Rev., 118, 1429–1443.

Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012a), Seasonal variation of CH4 emissions from central California. J. Geophys. Res., 117, D11306, doi:10.1029/2011JD016896.

Jeong, S., C. Zhao, A. E. Andrews, E. J. Dlugokencky, C. Sweeney, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012b), Seasonal variations in N_2O emissions from central California. Geophys. Res. Lett., 39, L16805, doi:10.1029/2012GL052307.

Kort, E. A., J. Eluszkiewicz, B. B. Stephens, J. B. Miller, C. Gerbig, T. Nehrkorn, B. C. Daube, J. O. Kaplan, S. Houweling, and S. C. Wofsy (2008), Emissions of CH_4 and N_2O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations. Geophy. Res. Lett., 35, L18808, doi:10.1029/2008GL034031.

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, C. A. Brainger, B. B. Stephens, P. S. Bakwin, and D. Y. Hollinger (2004), Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO_2 Budget and Rectification Airborne (COBRA) study. J. Geophys. Res., 109, D15304, doi:10.1029/2004JD004754.

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. J. Geophys. Res., 108(D16), 4493, doi:10.1029/2002JD003161.

McMillan, A. M. S., M. L. Goulden, and S. C. Tyler (2007), Stoichiometry of CH_4 and CO_2 flux in a California rice paddy, J. Geophys. Res., 112, G01008, doi:10.1029/2006JG000198.

Mellor G. L. and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems. Rev. Geophys. Space Phys. 20, 851–875.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2006), North American Regional Reanalysis. Bull. Amer. Meteor. Soc., 87 (3), 343-360.

Miller, S. M., et al. (2012), Regional sources of nitrous oxide over the United States: Seasonal variation and spatial distribution, J. Geophys. Res., 117, D06310, doi:10.1029/2011JD016951.

Montzka, S. A., E. J. Dlugokencky1, and J. H. Butler (2011) Non-CO_2 greenhouse gases and climate change. Nature, 476, 43-50.

Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled weather research and forecasting - stochastic time-inverted lagrangian transport (WRF-STILT) model. Meteor. Atmos. Phys., 107 (1), 51-64, doi:10.1007/s00703-010-0068-x.

Peischl, J., et al. (2012), Airborne observations of methane emissions from rice cultivation in the Sacramento Valley of California, J. Geophys. Res., 117, D00V25, doi:10.1029/2012JD017994.

Potter, C., S. Klooster, S. Hiatt, M. Fladeland, V. Genovese and P. Gross (2006), Methane Emissions from Natural Wetlands in the United States: Satellite-Derived Estimation Based on Ecosystem Carbon Cycling. Earth Interactions, 10, 1–12.

Press, W. H., S. A.Teukolsky, W. T. Vetterling, and B.P. Flannery (1992), Numerical Recipes, 2nd edition. Cambridge: Cambridge University Press.

Salas, W., P. Green, S. Frolking, C. Li and S. Boles (2006), Estimating Irrigation Water Use for California Agriculture: 1950s to Present. California Energy Commission, PIER Energy-Related Environmental Research. CEC-500-2006-057.

Skamarock, W.C., J. B. Klemp, J. Dudhia, D.O. Gill, D. M. Barker, X. Z. Huang, W. Wang, and J. G. Powers (2008), A description of the advanced research WRF version 3. Technical Note 475+STR. Mesoscale and Microscale Meteorology Division, NCAR, Boulder, Colorado.

Wennberg, P. O., W. Mui, D Wunch, E. A. Kort, D. R. Blake, E. L. Atlas, G. W. Santoni, S. C. Wofsy, G. S. Diskin, S. Jeong, and M. L. Fischer (2012), On the Sources of Methane to the Los Angeles Atmosphere. Environ. Sci. Technol., 46 (17), 9282 - 9289, doi:10.1021/es301138y.

Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California. J. Geophys. Res., 114, D16302, doi:10.1029/2008JD011671.